

A STATISTICAL APPROACH TO ZONAL POLYNOMIALS

TECHNICAL REPORT NO. 7

AKIMICHI TAKEMURA

JANUARY 1983

U. S. ARMY RESEARCH OFFICE  
CONTRACT DAAG29-82-K-0156

DEPARTMENT OF STATISTICS  
STANFORD UNIVERSITY  
STANFORD, CALIFORNIA



APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

A STATISTICAL APPROACH TO ZONAL POLYNOMIALS

TECHNICAL REPORT NO. 7

AKIMICHI TAKEMURA  
STANFORD UNIVERSITY

JANUARY 1983

U. S. ARMY RESEARCH OFFICE

CONTRACT DAAG 29-82-K-0156

DEPARTMENT OF STATISTICS  
STANFORD UNIVERSITY  
STANFORD, CALIFORNIA

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

THE VIEW, OPINIONS, AND/OR FINDINGS CONTAINED IN THIS REPORT ARE THOSE OF THE AUTHOR(S) AND SHOULD NOT BE CONSTRUED AS AN OFFICIAL DEPARTMENT OF THE ARMY POSITION, POLICY, OR DECISION, UNLESS SO DESIGNATED BY OTHER DOCUMENTATION.

## ACKNOWLEDGEMENTS

I wish to express my sincere thanks to my advisor Professor T.W. Anderson who represents one of the best traditions in the field of statistics, to Professor Persi Diaconis and Professor Ingram Olkin for reading this thesis and suggesting many improvements, and to Fred Huffer and Satish Iyengar for their careful and helpful comments. This research was supported in part by Office of Naval Research Contract N00014-75-C-0442 and U.S. Army Research Office Contract DAAG29-82-K-0156.

# Contents

	Page
Chapter 1. Introduction . . . . .	1
Chapter 2. Preliminaries on partitions and homogeneous symmetric polynomials. . . . .	5
2.1 Partitions. . . . .	5
2.2 Homogeneous symmetric polynomials. . . . .	8
Chapter 3. Derivation and some basic properties of zonal polynomials . . . . .	15
3.1 Definition of zonal polynomials. . . . .	15
3.2 Integral identities involving zonal polynomials . . . . .	23
3.3 An integral representation of zonal polynomials . . . . .	32
3.4 A generating function of zonal polynomials . . . . .	34
Chapter 4. More properties of zonal polynomials. . . . .	39
4.1 Majorization ordering . . . . .	39
4.2 Evaluation of ${}_1Y_p(I_k)$ . . . . .	44
4.3 More on integral identities . . . . .	47
4.4 Coefficients of $U_q$ in $Y_p$ . . . . .	52
4.4.1 Rank 1 and rank 2 cases. . . . .	52
4.4.2 Recurrence relations on the coefficients . . . . .	54
4.5 Conversion to coefficients of $M_q$ and James' partial differential equation . . . . .	59
4.5.1 Rank 1 and rank 2 cases. . . . .	60
4.5.2 Again on the generating function of zonal polynomials. . . . .	61
4.5.3 Recurrence relations of Section 4.4.2. . . . .	63
4.5.4 James' partial differential equation and recurrence relation. . . . .	64
4.6 Coefficients of $T_q$ in $Z_p$ . . . . .	67
4.7 Variations of the integral representation of zonal polynomials . . . . .	71
Chapter 5. Complex zonal polynomials. . . . .	75
5.1 The complex normal and the complex Wishart distributions.. . . .	75
5.2 Derivation and properties of complex zonal polynomials . . . . .	77
5.3 Schur functions and their determinantal expressions.. . . .	83
5.4 Relation between the real and the complex zonal polynomials. . . . .	89
Appendix: List of new results. . . . .	91
References . . . . .	92

## Introduction

In studying the behavior of power functions of various multivariate tests under alternative hypotheses it is essential to develop a distribution theory for noncentral distributions in the multivariate case. It is somewhat surprising that this is not straightforward. In fact even intuitively obvious results concerning the power functions of multivariate tests often require fairly elaborate arguments because of this difficulty. For a recent example of this see Olkin and Perlman (1980).

Let us take the noncentral  $\chi^2$  distribution and its multivariate analog, the noncentral Wishart distribution, as an example. The density of the noncentral  $\chi^2$  distribution is usually written as an infinite series. This series arises from the expansion of the exponential part of the normal density into a power series and its term by term integration with respect to irrelevant variables. In the multivariate case this integration becomes nontrivial involving an integration with respect to the Haar invariant measure on the orthogonal group. Anderson and Girshick (1944) and Anderson (1946) were the first systematic studies of this problem. James (1955a, 1955b) introduced the integration with respect to the Haar measure on the orthogonal group explicitly and made further progress. Herz (1955) developed a theory of hypergeometric functions in matrix arguments and expressed the density of the noncentral Wishart distribution by a hypergeometric function. Then James (1960, 1961) introduced "zonal polynomials" which have a special invariance property with respect to the Haar measure and the density of the noncentral Wishart distribution can be expressed in an infinite series involving zonal polynomials. This infinite series provides an explicit infinite

series expression for the hypergeometric function introduced by Herz. James (1960,1961b), Constantine (1963, 1966) and others worked out many other distributions and expressed them in a similar fashion. Thus zonal polynomials provide a unifying tool for the study of multivariate noncentral distributions.

In spite of this advantage zonal polynomials have been considered to be difficult to understand. This is mainly because the definition of zonal polynomials required rather extensive knowledge of group representation theory, spherical function theory, etc. This difficulty was hardly alleviated by more recent (abstract) treatments by Farrell(1976) and Kates(1980). Another reason is that explicit expressions for zonal polynomials are not known. Furthermore an infinite series involving zonal polynomials often converges very slowly. Because of these difficulties many noncentral multivariate densities which can be expressed in infinite series involving zonal polynomials have been rarely used and studied numerically.

The purpose of this paper is to present a self-contained development of zonal polynomials for people working with multivariate analysis. Our primary goal is to write a readable account of zonal polynomials in the framework of standard multivariate analysis and our development in the sequel might seem sometimes repetitive. Furthermore we pay fairly close attention to the numerical aspects of zonal polynomials, their coefficients, etc. No knowledge of extensive mathematics is needed but we assume that the reader is familiar with usual multivariate analysis (e.g. Anderson(1958)) and linear algebra. As is often the case with an elementary approach we have to do a good deal of computation, which sometimes becomes tedious. This is in a sense a necessary tradeoff in adopting an elementary approach. In any case usual material in multivariate analysis requires fairly heavy computation and the computation in the sequel does not seem too heavy.

Recently there have been several elementary treatments of zonal polynomials. Let us briefly discuss these approaches and their relations to the present work. Our starting point was Saw(1977) who derived many properties of zonal polynomials using basic properties of the multivariate normal and the Wishart distributions. Unfortunately at several points Saw(1977) refers to Constantine(1963) which in turn makes use of group representation theory. Therefore Saw(1977) is not entirely self-contained. Actually as Saw(1977) suggested, it turns out that only the elementary methods from Constantine(1963)

are needed to complete Saw's argument but it was not clear. Furthermore it seemed more advantageous to define zonal polynomials in a different way than Saw did and rearrange his logical steps, especially because his definition of zonal polynomials lacked a conceptual motivation. See Remark 3.12 on this point. In our approach zonal polynomials will be defined as eigenfunctions of an expectation or integral operator. Actually by considering the finite dimensional vector space of homogeneous symmetric polynomials of a given degree we work with vectors and matrices and we define zonal polynomials just as characteristic vectors of a certain matrix. This will be done in Section 3.1.

The idea of "eigenfunctions of expected value operators" is investigated in a more general framework in recent works by Kushner and Meisner(1980) and Kushner, Lebow, and Meisner(1981). In the second paper they gave a definition of zonal polynomials. They follow James' original idea but mostly use techniques of linear algebra and their approach is very helpful for the understanding of James' original definition. They consider the space of homogeneous polynomials of (elements of) a symmetric matrix variable  $\mathbf{A}$  whereas we consider the space of symmetric homogeneous polynomials of the characteristic roots of  $\mathbf{A}$ . In the former approach an extra step is needed to define zonal polynomials by requiring an orthogonal invariance. Therefore our approach in Section 3.1 seems to be more direct.

Another welcome addition to our limited literature of elementary treatment of zonal polynomials is provided by the recent book by Muirhead(1982), in which zonal polynomials are defined as eigenfunctions of a differential operator, or as solutions to partial differential equations. Although Muirhead proceeds informally and many points are illustrated rather than proved, the approach can be made precise. We discuss this in Section 4.5.4. From our prejudiced viewpoint our approach seems more natural, especially because the differential equation is hard to motivate.

None of these treatments seem to be as complete as the present work.

Zonal polynomials were originally defined using group representation theory and their properties were derived from the definition. An obvious alternative for us is to look at various properties satisfied by zonal polynomials and see whether these properties can be used as a definition. A definition must be first of all meaningful, namely it should define some mathematical object which exists in a unique way. It becomes a successful one if all



other properties can be derived from it. Is our definition successful? The answer seems to be affirmative. Actually we will derive almost all the known major properties of zonal polynomials in a self-contained way as well as many new results. Proofs are often new. The new contributions of the present work are listed in Appendix.

A good measure to judge our optimistic claim is a remarkable paper by James(1964). It is expository and contains many statements which seem to have never been proved in publication. In the sequel we will often refer to formula numbers in this paper.

We will not discuss various noncentral densities which are written as infinite series involving zonal polynomials. In addition to James(1964) we just refer to Subrahmaniam(1976) which provides extensive references for various applications of zonal polynomials. Once zonal polynomials are defined and their properties derived it seems not hard to write down those densities. We will not discuss hypergeometric functions of matrix arguments either.

On the other hand we included a development of complex zonal polynomials (zonal polynomials associated with the complex normal and the complex Wishart distributions). One reason for this is that our approach for the real case almost immediately carries over to the complex case. Another reason is that the existing theory of Schur functions provides explicit expressions for complex zonal polynomials. This becomes apparent if one compares Farrell(1980) and Macdonald(1979). See Chapter 5.

Finally let us describe subsequent chapters. Chapter 2 gives preliminary material on partitions and homogeneous symmetric polynomials. Definitions and notations should be checked since they vary from book to book. In Chapter 3 we define zonal polynomials and derive their major properties. If the reader is not much interested in computational aspects of zonal polynomials the material covered in Chapter 3 should suffice for usual applications. Chapter 4 generalizes and refines the results in Chapter 3. It deals largely with computation, coefficients, etc., of zonal polynomials. The development becomes inevitably more tedious. In Chapter 5 we apply our approach to complex zonal polynomials. We add some material from Macdonald(1979) and show that complex zonal polynomials are Schur functions.

# Preliminaries on partitions and homogeneous symmetric polynomials.

In this chapter we establish appropriate notations for partitions and homogeneous symmetric polynomials and summarize basic facts about them. A large part of the material in this section and section 5.3 is found in Macdonald(1979), Chapter 1.

## §2.1 Partitions.

A set of positive integers  $p = (p_1, \dots, p_\ell)$  is called a *partition* of  $n$  if  $n = p_1 + \dots + p_\ell$ . To denote  $p$  uniquely we order the elements as  $p_1 \geq p_2 \geq \dots \geq p_\ell$ .  $p_1, \dots, p_\ell$  are called *parts* of  $p$ ;  $\ell, p_1, n$  are

$$\begin{aligned} \ell &= \ell(p) = \text{length of } p = \text{number of parts,} \\ p_1 &= h(p) = \text{height of } p, \\ n &= |p| = \text{weight of } p. \end{aligned} \tag{2.1}$$

respectively. The multiplicity  $m_i$  of  $i$ , ( $i = 1, 2, \dots$ ) in  $p$  is defined as

$$m_i = \text{number of } j \text{ such that } p_j = i. \tag{2.2}$$

Using the  $m_i$ 's  $p$  is often denoted as  $p = (1^{m_1} 2^{m_2} \dots)$ . The set of all partitions of  $n$  is denoted by  $\mathcal{P}_n$  ( $= \{p : |p| = n\}$ ).

It is often convenient to look at  $p$  as having any number of additional zeros  $p = (p_1, \dots, p_\ell, 0, \dots, 0)$ . In this case it is understood that  $p_k = 0$  for  $k > \ell(p)$ . With this convention addition of two partitions is defined by  $(p + q)_i = p_i + q_i$ ,  $i = 1, 2, \dots$

A nice way of visualizing partitions is to associate the following diagrams to them. For  $p = (p_1, \dots, p_\ell)$  we associate a diagram which has  $p_i$  dots (or squares) in  $i$ -th row. For example the diagram of  $(4, 2, 2, 1)$  is given by

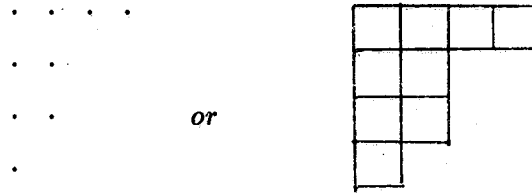


Figure 2.1.

We define the *conjugate partition*  $p'$  of  $p$  by means of this diagram, namely  $p'$  is a partition whose diagram is the transpose of the diagram of  $p$ . From Figure 2.1 we see  $(4, 2, 2, 1)' = (4, 3, 1, 1)$ . Clearly  $p'' = (p')' = p$ . Furthermore  $|p| = |p'|$ ,  $\ell(p) = h(p')$ ,  $h(p) = \ell(p')$ . More explicitly  $p'$  is determined by

$$(2.3) \quad m_i(p') = p_i - p_{i+1}, \quad i = 1, \dots, \ell.$$

Therefore for example

$$(2.4) \quad \begin{aligned} \ell(p') &= m_1(p') + m_2(p') + \dots \\ &= (p_1 - p_2) + (p_2 - p_3) + \dots \\ &= p_1 = h(p). \end{aligned}$$

Let  $s \geq h(p), t \geq \ell(p)$ . We define

$$(2.5) \quad p_{s,t}^* = (s - p_t, s - p_{t-1}, \dots, s - p_1)$$

From Figure 2.2 we have  $(4, 2, 2, 1)_{4,5}^* = (4, 3, 2, 2, 0)$ . Note that

$$(2.6) \quad |p_{s,t}^*| = st - |p|$$

$s$

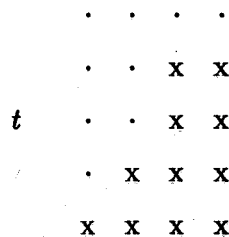


Figure 2.2.

Now we introduce two orderings in  $\mathcal{P}_n$ . The first one is called the *lexicographic ordering* ( $>$ ). In this ordering  $p$  is said to be *higher than*  $q$  ( $p > q$ ) if

$$(2.7) \quad p_1 = q_1, \dots, p_{k-1} = q_{k-1}, p_k > q_k \quad \text{for some } k.$$

This is a total ordering. For example  $\mathcal{P}_4$  is ordered as  $(4) > (3,1) > (2,2) > (2,1,1) > (1,1,1,1)$ .

This ordering is preserved by addition.

**Lemma 2.1.** *If  $p^1 \geq q^1, p^2 \geq q^2$  then  $p^1 + p^2 \geq q^1 + q^2$  with equality iff  $p^1 = p^2, q^1 = q^2$ .*

*Proof:* Let  $k_i, i = 1, 2$ , be the first index such that  $p_{k_i}^i > q_{k_i}^i$ . ( $k_i$  is defined to be  $\infty$  if  $p^i = q^i$ .) Let  $k = \min(k_1, k_2)$ . Then

$$(2.8) \quad \begin{aligned} p_j^1 + p_j^2 &= q_j^1 + q_j^2, & \text{for } j < k, \\ p_k^1 + p_k^2 &> q_k^1 + q_k^2. \end{aligned}$$

Hence  $p^1 + p^2 \geq q^1 + q^2$ . Equality holds if and only if  $k = \infty$  or  $p^1 = q^1, p^2 = q^2$ . ■

Another ordering is the *majorization ordering*.  $p$  *majorizes*  $q$  ( $p \succ q$ ) if and only if

$$(2.9) \quad p_1 \geq q_1, p_1 + p_2 \geq q_1 + q_2, \dots, p_1 + \dots + p_k \geq q_1 + \dots + q_k, \dots$$

Note that for  $k \geq \max(\ell(p), \ell(q))$  the equality holds because both sides are equal to the weight  $n$ . Majorization is a partial ordering and it is stronger than the lexicographic ordering:

**Lemma 2.2.** *If  $p \succ q$  then  $p \geq q$ .*

*Proof:* Suppose  $p_1 = q_1, \dots, p_{k-1} = q_{k-1}, p_k \neq q_k$ . Then  $p_1 + \dots + p_k \geq q_1 + \dots + q_k$  implies  $p_k > q_k$ . Hence  $p > q$ . ■

**Remark 2.1.** *The converse of Lemma 2.2 is false. For example  $(3,1,1,1) > (2,2,2)$  but there is no majorization between these two.*

Analogous to Lemma 2.1 we have

**Lemma 2.3.** *If  $p^1 \succ q^1, p^2 \succ q^2$ , then  $p^1 + p^2 \succ q^1 + q^2$  with equality iff  $p^1 = p^2, q^1 = q^2$ .*

*Proof:* For any  $k$

$$(p_1^1 + p_1^2) + \cdots + (p_k^1 + p_k^2) \geq (q_1^1 + q_1^2) + \cdots + (q_k^1 + q_k^2)$$

with equality iff  $p_1^i + \cdots + p_k^i = q_1^i + \cdots + q_k^i, i = 1, 2.$  ■

## §2.2 Homogeneous symmetric polynomials.

Let  $f(x_1, \dots, x_k)$  be a polynomial in  $x_1, \dots, x_k$ .  $f$  is *homogeneous (of degree  $n$ )* if  $f$  has only  $n$ -th degree terms.  $f$  is *symmetric* if

$$(2.10) \quad f(x_1, \dots, x_k) = f(x_{i_1}, \dots, x_{i_k}),$$

where  $(i_1, \dots, i_k)$  is any permutation of  $(x_1, \dots, x_k)$ . Let  $V_n$  denote the set of all  $n$ -th degree homogeneous symmetric polynomials including the constant  $f \equiv 0$ . We look at  $V_n$  as a vector space where addition is the usual addition of polynomials. Let  $f \in V_n$  and suppose that  $f$  has a term  $ax_1^{p_1} \cdots x_\ell^{p_\ell}$   $((p_1, \dots, p_\ell) \in \mathcal{P}_n)$ , then by symmetry it also has a term  $ax_{i_1}^{p_1} \cdots x_{i_\ell}^{p_\ell}$  where  $i_1, \dots, i_\ell$  are distinct integers taken from  $(1, \dots, k)$ . Counting all different terms we see that  $f$  can be written as a linear combination of *monomial symmetric functions*  $M_p, p \in \mathcal{P}_n$ ,

$$(2.11) \quad f = \sum_{p \in \mathcal{P}_n} a_p M_p,$$

where

$$(2.12) \quad M_p = \sum_{(i_1, \dots, i_\ell) \subset (1, \dots, k)} x_{i_1}^{p_1} \cdots x_{i_\ell}^{p_\ell}.$$

In (2.12) we count only distinguishable terms. For example

$$(2.13) \quad M_{(1,1)} = \sum_{i < j} x_i x_j.$$

Sometimes it is more convenient to use *augmented monomial symmetric function*  $\mathcal{A}M_p$  for which the summation in (2.12) is over all permutations of  $\ell$  different integers from  $(1, \dots, k)$ . Therefore

$$(2.14) \quad \mathcal{A}M_{(1,1)} = \sum_{i \neq j} x_i x_j = 2M_{(1,1)}.$$

In general

$$(2.15) \quad \mathcal{A}M_p = \left( \prod_{i=1}^{h(p)} m_i! \right) M_p.$$

where  $(p_1, \dots, p_\ell) = (1^{m_1} 2^{m_2} \dots)$ .

We note that in (2.11) the number of variables  $k$  does not play an explicit role. Actually  $M_p$  can be defined for any number of variables by (2.12) and

$$(2.16) \quad M_p(x_1, \dots, x_k, 0, \dots, 0) = M_p(x_1, \dots, x_k).$$

Hence it suffices to consider  $M_p$  which is defined for sufficiently large number of variables. Now suppose

$$(2.17) \quad \sum_{p \in \mathcal{P}_n} a_p M_p = 0,$$

We look at terms of the form  $x_1^{q_1} \dots x_\ell^{q_\ell}$ . Differentiating (2.17)  $p_i$  times with respect to  $x_i$ ,  $i=1, \dots, \ell$  we have  $a_p = 0$ . Hence  $M_p, p \in \mathcal{P}_n$  are linearly independent in  $V_n$ . (Of course if  $k < \ell(p)$  then  $M_p(x_1, \dots, x_k) = 0$  which is linearly dependent in a trivial sense. But as above we consider  $k$  to be sufficiently large. For more detail see Section 4.1.) From (2.11) and (2.17) it follows that  $\{M_p, p \in \mathcal{P}_n\}$  forms a basis of  $V_n$ . This is a rather obvious basis.

We want to consider other bases. The following lemma is useful for this purpose.

**Lemma 2.4.** *If  $A$  is an upper triangular matrix with nonzero diagonal elements, then  $A^{-1}$  has the same property.*

*Proof:* By assumption  $|A| \neq 0$ . Hence  $A^{-1}$  exists. Let  $A^{-1} = \{a^{ij}\}$ . Then  $a^{ij} = \Delta_{ji}/|A|$ , where  $\Delta_{ji}$  is a cofactor of  $A$ . By triangularity of  $A$ ,  $\Delta_{ji} = 0$  for  $i > j$ . ■

**Corollary 2.1.** *If  $A$  is an upper triangular matrix with diagonal elements 1 and integral offdiagonal elements, then  $A^{-1}$  has the same property.*

*Proof:*  $|A| = 1$ . Furthermore  $\Delta_{ij}$  is an integer since the elements of  $A$  are integers. ■

Now we consider products of elementary symmetric functions. Let

$$(2.18) \quad u_r = \sum_{i_1 < \dots < i_r} x_{i_1} \cdots x_{i_r}$$

be the  $r$ -th elementary symmetric function. For  $p \in \mathcal{P}_n$  we define

$$(2.19) \quad \mathcal{U}_p = u_1^{p_1-p_2} u_2^{p_2-p_3} \cdots u_\ell^{p_\ell}.$$

The degree of  $\mathcal{U}_p$  is

$$(2.20) \quad (p_1 - p_2) + 2(p_2 - p_3) + \cdots + \ell p_\ell = p_1 + \cdots + p_\ell = n.$$

Hence  $\mathcal{U}_p \in V_n$ .  $\mathcal{U}_p$  defined by (2.19) corresponds to  $\mathcal{U}_p$  in Macdonald's notation (1979).

**Lemma 2.5.**

$$(2.21) \quad \mathcal{U}_p = \mathcal{M}_p + \sum_{q < p} a_{pq} \mathcal{M}_q,$$

where  $a_{pq}$  are integers.

*Proof:* Consider monomial terms of the form  $x_1^{q_1} x_2^{q_2} \cdots x_\ell^{q_\ell}$ ,  $q = (q_1, \dots, q_\ell) \in \mathcal{P}_n$ . Now

$$\mathcal{U}_p = (x_1 + \cdots)^{p_1-p_2} (x_1 x_2 + \cdots)^{p_2-p_3} \cdots (x_1 \cdots x_\ell + \cdots)^{p_\ell}.$$

Hence the highest order term obtained by expanding  $\mathcal{U}_p$  is

$$x_1^{p_1-p_2} (x_1 x_2)^{p_2-p_3} \cdots (x_1 \cdots x_\ell)^{p_\ell} = x_1^{p_1} x_2^{p_2} \cdots x_\ell^{p_\ell},$$

which has coefficient 1. It is clear that other terms are lower in the lexicographic ordering and have integral coefficients. ■

**Remark 2.2.** For a stronger result see Lemma 4.1.

We order  $\mathcal{M}_p, \mathcal{U}_p, p \in \mathcal{P}_n$  according to the lexicographic ordering and form two vectors:

$$(2.22) \quad \mathcal{M} = \begin{pmatrix} \mathcal{M}_{(n)} \\ \mathcal{M}_{(n-1,1)} \\ \cdot \\ \cdot \\ \cdot \\ \mathcal{M}_{(1^n)} \end{pmatrix}, \quad \mathcal{U} = \begin{pmatrix} \mathcal{U}_{(n)} \\ \mathcal{U}_{(n-1,1)} \\ \cdot \\ \cdot \\ \cdot \\ \mathcal{U}_{(1^n)} \end{pmatrix}.$$

Then Lemma 2.5 implies that

$$(2.23) \quad \mathcal{U} = A\mathcal{M}, \quad A = (a_{pq}),$$

where  $A$  is a matrix satisfying the condition of Corollary 2.1. Therefore considering  $A^{-1} = (a^{pq})$  we obtain

$$(2.24) \quad \mathcal{M}_p = \mathcal{U}_p + \sum_{q < p} a^{pq} \mathcal{U}_q,$$

where  $a^{pq}$  are integers. We see that  $\{\mathcal{U}_p, p \in \mathcal{P}_n\}$  forms another basis of  $V_n$ .

Product of  $\mathcal{U}$  functions corresponds to the addition of partitions.

**Lemma 2.6.**

$$(2.25) \quad \mathcal{U}_p \mathcal{U}_q = \mathcal{U}_{p+q}.$$

*Proof:*

$$\begin{aligned} \mathcal{U}_p \mathcal{U}_q &= u_1^{p_1-p_2} u_2^{p_2-p_3} \dots u_1^{q_1-q_2} u_2^{q_2-q_3} \dots \\ &= u_1^{(p_1+q_1)-(p_2+q_2)} u_2^{(p_2+q_2)-(p_3+q_3)} \dots \\ &= \mathcal{U}_{p+q}. \end{aligned}$$

■



The third basis of  $V_n$  is given by product of power sums. Let

$$(2.26) \quad t_r = \sum x_i^r.$$

For  $p \in \mathcal{P}_n$  we define

$$(2.27) \quad \tau_p = t_1^{p_1 - p_2} t_2^{p_2 - p_3} \dots t_\ell^{p_\ell}.$$

$\tau_p$  defined by (2.27) corresponds to  $\tau_p$  in Macdonald(1979) and in Saw(1977). Here we prefer the above definition because of the simpler relation between  $\mathcal{U}_p$  and  $\tau_p$ .

Let

$$(2.28) \quad U(s) = \prod (1 + sx_i) = 1 + u_1 s + u_2 s^2 + \dots$$

be a generating function of  $u$ 's. Then

$$(2.29) \quad \begin{aligned} \log U(s) &= \sum \log(1 + sx_i) \\ &= st_1 - \frac{s^2}{2} t_2 + \dots + (-1)^{r-1} \frac{s^r}{r} t_r + \dots \end{aligned}$$

On the other hand

$$(2.30) \quad \log U(s) = (u_1 s + u_2 s^2 + \dots) - \frac{1}{2}(u_1 s + u_2 s^2 + \dots)^2 + \dots$$

Comparing coefficients of  $s^r$  in (2.29) and (2.30) we see

$$(2.31) \quad \begin{aligned} t_r &= (-1)^{r-1} r \{u_r + \sum_{q > (1^r), q \in \mathcal{P}_r} a_{rq} \mathcal{U}_q\} \\ &= (-1)^{r-1} r \{\mathcal{U}_{(1^r)} + \sum_{q > (1^r), q \in \mathcal{P}_r} a_{rq} \mathcal{U}_q\}. \end{aligned}$$

It is straightforward to show that

$$(2.32) \quad \begin{aligned} a_{rq} &= \frac{(-1)^{q_1-1}}{q_1} \binom{q_1}{q_1 - q_2, q_2 - q_3, \dots, q_{\ell(q)}} \\ &= \frac{(-1)^{q_1-1} (q_1 - 1)!}{(q_1 - q_2)! \dots q_{\ell(q)}!}. \end{aligned}$$

Now

$$\begin{aligned}
 \tau_p &= \prod_{r=1}^{\ell(p)} t_r^{p_r - p_{r+1}} \\
 (2.33) \quad &= \prod_{r=1}^{\ell(p)} \left[ (-1)^{r-1} r \{ \mathcal{U}_{(1^r)} + \sum_{q > (1^r), q \in \mathcal{P}_r} a_{rq} \mathcal{U}_q \} \right]^{p_r - p_{r+1}}.
 \end{aligned}$$

By Lemma 2.1 and Lemma 2.6 the lowest order term in (2.33) is given by

$$\begin{aligned}
 &\prod_{r=1}^{\ell} [(-1)^{r-1} r \mathcal{U}_{(1^r)}]^{p_r - p_{r+1}} \\
 (2.34) \quad &= \prod_{r=1}^{\ell} [(-1)^{r-1} r]^{p_r - p_{r+1}} u_1^{p_1 - p_2} u_2^{p_2 - p_3} \dots u_{\ell}^{p_{\ell}} \\
 &= (-1)^{|p| - p_1} \left( \prod_{r=1}^{\ell} r^{p_r - p_{r+1}} \right) \mathcal{U}_p.
 \end{aligned}$$

Hence

**Lemma 2.7.**

$$(2.35) \quad \tau_p = \sum_{q \geq p} a_{pq} \mathcal{U}_q,$$

where

$$(2.36) \quad a_{pp} = (-1)^{|p| - p_1} \prod_{r=1}^{\ell(p)} r^{p_r - p_{r+1}} \neq 0.$$

Let

$$\tau = \begin{pmatrix} \tau_{(n)} \\ \tau_{(n-1,1)} \\ \cdot \\ \cdot \\ \cdot \\ \tau_{(1^n)} \end{pmatrix}.$$

Then Lemma 2.7 shows that

$$(2.37) \quad \tau = F \mathcal{U},$$

where  $F$  is lower triangular with nonzero diagonal elements. Hence  $\{ \tau_p, p \in \mathcal{P}_n \}$  forms a basis of  $V_n$ .

**Remark 2.3.** *To show that  $\{\tau_p, p \in \mathcal{P}_n\}$  is a basis it is much easier to note*

$$\tau_{p'} = \mathcal{A}M_p + \sum_{q > p} a_{pq} \mathcal{A}M_q,$$

*where  $a_{pq}$  are integers and use Corollary 2.1. But we will use Lemma 2.7 in Section 4.6.*

We study symmetric functions further in Section 5.3. However the material covered so far suffices to derive zonal polynomials which form another basis of  $V_n$ .

## Derivation and some basic properties of zonal polynomials

In this chapter we define (real) zonal polynomials and derive some properties. Some remarks on notation seem appropriate here. We define zonal polynomials as characteristic vectors of a certain linear transformation  $\tau$  from  $V_n$  to  $V_n$ . The normalization is rather arbitrary for a characteristic vector and many properties of zonal polynomials are independent of particular normalization. Corresponding to different normalizations, different symbols such as  $Z_p, C_p$  have been used to denote zonal polynomials. We find it advantageous to use still another normalization in addition to those corresponding to  $Z_p, C_p$ . Considering these circumstances we use  $y_p$  for an unnormalized zonal polynomial.  ${}_1y_p$  is used to denote a zonal polynomial normalized so that the coefficient of  $U_p$  or  $M_p$  is 1.

### §3.1 Definition of zonal polynomials

As mentioned earlier we define zonal polynomials as characteristic vectors of a certain matrix. The matrix in question will be triangular and we begin by a lemma concerning a triangular matrix and its characteristic vectors.

**Lemma 3.1.** *Let  $T = (t_{ij})$  be an  $n \times n$  upper triangular matrix with distinct diagonal elements. Let  $A = \text{diag}(t_{11}, \dots, t_{nn})$ . Then there exists an upper triangular matrix  $E$  satisfying*

$$(3.1) \quad ET = AE.$$

$\Xi$  is uniquely determined up to a (possibly different) multiplicative constant for each row.

*Proof:* The diagonal elements of  $T$  are the characteristic roots. (3.1) shows that the  $i$ -th row  $\xi'_i$  of  $\Xi$  is a characteristic vector (from the left) associated with  $t_{ii}$ . Let

$$(3.2) \quad y'T = (y_1, y'_2) \begin{pmatrix} t_{11} & t'_1 \\ 0 & T_2 \end{pmatrix} = t_{11}y'.$$

Given  $y_1 = c$  (3.2) becomes

$$(3.3) \quad (t_{11}c, ct'_1 + y'_2T_2) = (t_{11}c, t_{11}y'_2).$$

Therefore

$$(3.4) \quad y'_2(T_2 - t_{11}I) = -ct'_1.$$

By assumption  $T_2 - t_{11}I$  is an upper triangular matrix with nonzero diagonal elements.

Hence  $y'_2$  is uniquely determined as

$$(3.5) \quad y'_2 = -ct'_1(T_2 - t_{11}I)^{-1} \quad (c \neq 0).$$

Now for  $i > 1$  let

$$(3.6) \quad y'T = t_{ii}y'.$$

Then considering first  $i - 1$  elements we have

$$(3.7) \quad \begin{aligned} y_1t_{11} &= t_{ii}y_1 \\ y_1t_{12} + y_2t_{22} &= t_{ii}y_2 \\ &\dots \\ y_1t_{1,i-1} + \dots + y_{i-1}t_{i-1,i-1} &= t_{ii}y_{i-1}. \end{aligned}$$

Inductively solving this we obtain  $0 = y_1 = \dots = y_{i-1}$ . Then (3.6) reduces to an equation like (3.2) with a smaller dimensionality. Hence arguing as above we see that given  $y_i = c$ ,  $(y_{i+1}, \dots, y_n)$  is uniquely determined. ■

**Remark 3.1.** *This lemma seems to be well known to people in numerical analysis although an explicit reference is not easy to find. It is very briefly mentioned on page 365 of Stewart(1973) in connection with the QR algorithm. The QR algorithm is designed to transform a general matrix to a triangular form in order to obtain the characteristic roots and vectors.*

For a  $k \times k$  matrix  $A = (a_{ij})$  we denote its characteristic roots by

$$(3.8) \quad \alpha = (\alpha_1, \dots, \alpha_k) = \lambda(A),$$

and (the determinant of) a principal minor by

$$(3.9) \quad A(i_1, \dots, i_\ell) = \begin{vmatrix} a_{i_1 i_1} & \dots & a_{i_1 i_\ell} \\ \vdots & & \vdots \\ a_{i_\ell i_1} & \dots & a_{i_\ell i_\ell} \end{vmatrix}.$$

For a matrix argument we define

$$\mathcal{U}_p(A) = \mathcal{U}_p(\alpha) = \mathcal{U}_p(\lambda(A)).$$

As is easily seen by expanding the determinant  $|A - \lambda I|$  the  $r$ -th elementary symmetric function of the roots of a matrix  $A$  is equal to the sum of  $r \times r$  principal minors, namely

$$(3.10) \quad \mathcal{U}_{(1^r)}(A) = u_r(\alpha_1, \dots, \alpha_k) = \sum_{i_1 < \dots < i_r} A(i_1, \dots, i_r).$$

(See Theorem 7.1.2 of Mirsky(1955) for example.) Hence

$$(3.11) \quad \mathcal{U}_p(A) = \left\{ \sum_{i_1} A(i_1) \right\}^{p_1 - p_2} \left\{ \sum_{i_1 < i_2} A(i_1, i_2) \right\}^{p_2 - p_3} \dots$$

Accordingly  $V_n(A)$  denotes the vector space of  $n$ -th degree homogeneous symmetric polynomials in the roots of  $A$ .

Now consider a (linear) transformation  $\tau_\nu : V_n(A) \rightarrow V_n(A)$  defined by

$$(3.12) \quad (\tau_\nu(\mathcal{U}_p))(A) = (\tau_\nu \mathcal{U}_p)(A) = \mathcal{E}_W \{ \mathcal{U}_p(AW) \},$$

where  $A$  is a symmetric matrix and  $W$  is a random symmetric matrix having a Wishart distribution  $\mathcal{W}(I_k, \nu)$ . Here  $\mathcal{W}(\Sigma, \nu)$  denotes the Wishart distribution with covariance  $\Sigma$  and degrees of freedom  $\nu$ . ( $\tau_\nu$  is defined for the basis  $\{\mathcal{U}_p\}$  by (3.12) and for general elements of  $V_n(A)$   $\tau_\nu$  is given by the linearity of expectation.) First we need to verify:

**Lemma 3.2.** *If  $A$  is symmetric, then  $(\tau_\nu \mathcal{U}_p)(A) \in V_n(A)$ .*

*Proof:* Since  $A$  is symmetric it can be written as  $A = \Gamma D \Gamma'$  where  $\Gamma$  is orthogonal and  $D = \text{diag}(\alpha_1, \dots, \alpha_k)$ . Now  $\mathcal{U}_p(AW) = \mathcal{U}_p(\Gamma D \Gamma' W) = \mathcal{U}_p(D \Gamma' W \Gamma)$  because the nonzero roots are invariant when the matrices are permuted cyclically. Since the distribution of  $\Gamma' W \Gamma$  is the same as the distribution of  $W$ , we can take  $A = \text{diag}(\alpha_1, \dots, \alpha_k)$  without loss of generality. Then

$$(3.13) \quad AW(i_1, \dots, i_r) = (\alpha_{i_1} \cdots \alpha_{i_r}) W(i_1, \dots, i_r).$$

For example

$$(3.14) \quad AW(1, 2) = \begin{vmatrix} \alpha_1 w_{11} & \alpha_1 w_{12} \\ \alpha_2 w_{21} & \alpha_2 w_{22} \end{vmatrix} = \alpha_1 \alpha_2 \begin{vmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{vmatrix}.$$

From (3.10) and (3.13) the  $r$ -th elementary symmetric function of the characteristic roots of  $AW$  can be written as

$$(3.15) \quad u_r(\lambda(AW)) = \sum_{i_1 < \dots < i_r} \alpha_{i_1} \cdots \alpha_{i_r} W(i_1, \dots, i_r).$$

Substituting this into (3.11) and taking the expectation we obtain

$$(3.16) \quad (\tau_\nu \mathcal{U}_p)(A) = \mathcal{E}_W \left( \sum_{i_1} \alpha_{i_1} W(i_1) \right)^{p_1 - p_2} \left( \sum_{i_1 < i_2} \alpha_{i_1} \alpha_{i_2} W(i_1, i_2) \right)^{p_2 - p_3} \dots$$

Clearly this belongs to  $V_n(A)$ . ■

$\tau_\nu$  has the following triangular property.

**Corollary 3.1.**

$$(3.17) \quad (\tau_\nu \mathcal{U}_p)(A) = \lambda_{\nu p} \mathcal{U}_p(A) + \sum_{q < p} a_{pq} \mathcal{U}_q(A).$$

*Proof:* As in the proof of Lemma 2.5 the highest monomial term in (3.16) is of the form

$$(3.18) \quad \alpha_1^{p_1} \alpha_2^{p_2} \cdots \alpha_\ell^{p_\ell} \mathcal{E}_W \{ W(1)^{p_1 - p_2} W(1, 2)^{p_2 - p_3} \cdots W(1, \dots, \ell)^{p_\ell} \}.$$

Then using (2.24) we see that  $(\tau_\nu \mathcal{U}_p)(A)$  expressed as a linear combination of  $\mathcal{U}_q$ 's involves only  $q$ 's such that  $q \leq p$ . In particular the leading coefficient is

$$(3.19) \quad \lambda_{\nu p} = \mathcal{E}_W \{ W(1)^{p_1 - p_2} W(1, 2)^{p_2 - p_3} \cdots W(1, \dots, \ell)^{p_\ell} \}.$$

■

**Remark 3.2.** The constants  $a_{pq}$  in (3.17) depend on the degrees of freedom  $\nu$ .

**Remark 3.3.** To be complete we have to verify that (3.12) does not depend on the number of variables  $k$  or more precisely we need to verify

$$(3.20) \quad (\tau_\nu \mathcal{U}_p)(\alpha_1, \dots, \alpha_k, 0, \dots, 0) = (\tau_\nu \mathcal{U}_p)(\alpha_1, \dots, \alpha_k),$$

for any number  $(m)$  of additional zeros. Note that the left hand side is defined using expectation with respect to  $\mathcal{W}(I_{k+m}, \nu)$ . Now recall that the marginal distribution of the  $k \times k$  upper left hand corner of  $\mathcal{W}(I_{k+m}, \nu)$  is  $\mathcal{W}(I_k, \nu)$  and (3.16) depends only on the  $k \times k$  upper left hand corner of the Wishart matrix. Therefore (3.20) holds.

By Corollary 3.1  $\tau_\nu$  expressed in an appropriate matrix form is an upper triangular matrix. In order to apply Lemma 3.1 we want to evaluate the "diagonal elements"  $\lambda_{\nu p}$  in (3.19). For that purpose we use the following well known result.

**Lemma 3.3.** Let  $W$  be distributed according to  $\mathcal{W}(I_k, \nu)$ . Let  $T = (t_{ij})$  be a lower triangular matrix with nonnegative diagonal elements such that  $W = TT'$ . Then  $t_{ij}$ ,  $i \geq j$ , are independently distributed as  $t_{ij} \sim \mathcal{N}(0, 1)$ ,  $i > j$ ,  $t_{ii} \sim \chi(\nu - i + 1)$  where  $\chi(\nu - i + 1)$  denotes the chi-distribution with  $\nu - i + 1$  degrees of freedom.

*Proof:* This is implicit in the derivation of Wishart distribution in Anderson (1958), Chapter 7, or see Srivastava and Khatri(1979), Corollary 3.2.4. ■

**Corollary 3.2.**

$$(3.21) \quad \begin{aligned} \lambda_{\nu p} &= 2^n \prod_{i=1}^{\ell} \Gamma[p_i + \frac{1}{2}(\nu + 1 - i)] / \Gamma[\frac{1}{2}(\nu + 1 - i)] \\ &= 2^n \prod_{i=1}^{\ell} \left( \frac{\nu + 1 - i}{2} \right)_{p_i} \\ &= \nu(\nu + 2) \cdots (\nu + 2(p_1 - 1)) \\ &\quad \cdot (\nu - 1)(\nu + 1) \cdots (\nu - 1 + 2(p_2 - 1)) \\ &\quad \cdots \\ &\quad \cdot (\nu - \ell + 1) \cdots (\nu - \ell + 1 + 2(p_\ell - 1)), \end{aligned}$$



where  $\ell = \ell(p)$  and  $(a)_k = a(a+1)\cdots(a+k-1)$ .

*Proof:* Note

$$(3.22) \quad W(1, \dots, r) = (t_{11} \cdots t_{rr})^2.$$

Substituting this into (3.19) we obtain

$$(3.23) \quad \lambda_{\nu p} = \mathcal{E} \{ t_{11}^{2p_1} t_{22}^{2p_2} \cdots t_{\ell\ell}^{2p_\ell} \}.$$

Now  $t_{ii}^2$  is distributed according to  $\chi^2(\nu - i + 1)$  and  $\mathcal{E} t_{ii}^{2p_i} = (\nu - i + 1)(\nu - i + 3) \cdots (\nu - i + 1 + 2(p_i - 1))$ . From this we obtain (3.21). ■

This proof is given in Constantine(1963) in a slightly different form.

Using the vector notation introduced in (2.22) let

$$(3.24) \quad \tau_\nu(\mathcal{U}) = \begin{pmatrix} \tau_\nu(\mathcal{U}_{(n)}) \\ \tau_\nu(\mathcal{U}_{(n-1,1)}) \\ \vdots \\ \tau_\nu(\mathcal{U}_{(1^n)}) \end{pmatrix}.$$

Then Corollary 3.1 shows that

$$(3.25) \quad \tau_\nu(\mathcal{U}) = T_\nu \mathcal{U},$$

where  $T_\nu$  is an upper triangular matrix with diagonal elements  $t_{pp} = \lambda_{\nu p}$ .  $T_\nu$  almost fits the condition of Lemma 3.1. The question now is what  $\nu$  to take. Actually a particular choice of  $\nu$  does not matter; we have:

**Lemma 3.4.** *There exists an upper triangular matrix  $\Xi$  such that*

$$(3.26) \quad \Xi T_\nu = \Lambda_\nu \Xi, \quad \text{for all } \nu,$$

where  $\Lambda_\nu = \text{diag}(\lambda_{\nu p}, p \in \mathcal{P}_n)$ .  $\Xi$  is uniquely determined up to a (possibly different) multiplicative constant for each row.

Lemma 3.4 shows that  $T_\nu$  has the same set of characteristic vectors (from the left) for all  $\nu$ . A proof of this will be given later in this section. Now we define zonal polynomials using this  $\Xi$ .

**Definition 3.1.** (*zonal polynomials*)

Let  $\Xi$  be as in Lemma 3.4. Zonal polynomials  $y_p, p \in \mathcal{P}_n$ , are defined by

$$(3.27) \quad y = \begin{pmatrix} y_{(n)} \\ y_{(n-1,1)} \\ \vdots \\ y_{(1^n)} \end{pmatrix} = \Xi u.$$

**Remark 3.4.**  $\Xi$  is upper triangular and therefore  $y_p$  is a linear combination of  $u_q$ 's (or  $M_q$ 's) for which  $q \leq p$ . It follows that  $\{y_p, p \in \mathcal{P}_n\}$  forms a basis of  $V_n$ .

**Remark 3.5.** Since each row of  $\Xi$  is determined uniquely up to a multiplicative constant  $y_p$  is determined up to normalization. We use  $y_p$  to denote an unnormalized zonal polynomial.

In order to prove Lemma 3.4 we first establish that the  $T_\nu$ 's commute with each other.

**Lemma 3.5.**

$$(3.28) \quad T_\nu T_\mu = T_\mu T_\nu.$$

*Proof:* For a symmetric positive semi-definite matrix  $A$  let  $A^{\frac{1}{2}}$  be the symmetric positive semi-definite square root, i.e.,  $A^{\frac{1}{2}} = \Gamma D^{\frac{1}{2}} \Gamma'$  where  $\Gamma$  is orthogonal and  $D$  is diagonal in  $A = \Gamma D \Gamma'$ . Now let  $W, V$  be independently distributed according to  $\mathcal{W}(I_k, \nu)$ ,  $\mathcal{W}(I_k, \mu)$  respectively. Consider

$$(3.29) \quad \mathcal{E}_{W,V} \{u(A^{\frac{1}{2}} V A^{\frac{1}{2}} W)\},$$

where  $u = (u_{(n)}, u_{(n-1,1)}, \dots, u_{(1^n)})'$ . Taking expectation with respect to  $W$  first we obtain

$$(3.30) \quad \begin{aligned} & \mathcal{E}_{W,V} \{u(A^{\frac{1}{2}} V A^{\frac{1}{2}} W)\} \\ &= \mathcal{E}_V \{T_\nu u(A^{\frac{1}{2}} V A^{\frac{1}{2}})\} \\ &= \mathcal{E}_V \{T_\nu u(AV)\} \\ &= T_\nu \mathcal{E}_V \{u(AV)\} \\ &= T_\nu T_\mu u(A). \end{aligned}$$

We used the cyclic permutation of the matrices since nonzero characteristic roots are invariant. Similarly taking expectation with respect to  $V$  first we obtain

$$(3.31) \quad \mathcal{E}_{W,V} \{ \mathcal{U}(A^{\frac{1}{2}} V A^{\frac{1}{2}} W) \} = T_{\mu} T_{\nu} \mathcal{U}(A).$$

Hence  $T_{\nu} T_{\mu} \mathcal{U}(A) = T_{\mu} T_{\nu} \mathcal{U}(A)$  for any symmetric  $A$ . Therefore  $T_{\nu} T_{\mu} = T_{\mu} T_{\nu}$ . ■

See Theorem 2.2 of Kushner, Lebow, and Meisner(1981) for an analogous result in a more general framework.

Now we give a proof of Lemma 3.4.

*Proof of Lemma 3.4:* Consider  $\lambda_{\nu p}$  given by (3.21). Let us look at  $\lambda_{\nu p}$  as a polynomial in  $\nu$ . They are different polynomials since they have different sets of roots. Now two different polynomials can match only finite number of times. It follows that for sufficiently large  $\nu_0$ ,  $\lambda_{\nu_0 p}$ ,  $p \in \mathcal{P}_n$ , are all different. Let  $\nu_0$  be fixed such that  $\lambda_{\nu_0 p}$ ,  $p \in \mathcal{P}_n$  are all different. Let  $\mathcal{E}$  be the matrix in (3.1) with  $T$  replaced by  $T_{\nu_0}$ . Note that the uniqueness part of Lemma 3.4 is already established now. Let  $A = \text{diag}(\lambda_{\nu_0 p}, p \in \mathcal{P}_n)$ . Then for any  $\mu$   $A(\mathcal{E}T_{\mu}) = (A\mathcal{E})T_{\mu} = (\mathcal{E}T_{\nu_0})T_{\mu} = \mathcal{E}(T_{\nu_0}T_{\mu}) = \mathcal{E}(T_{\mu}T_{\nu_0}) = (\mathcal{E}T_{\mu})T_{\nu_0}$ , or  $A\mathcal{E}_1 = \mathcal{E}_1T_{\nu_0}$  where  $\mathcal{E}_1 = \mathcal{E}T_{\mu}$ . Now by the uniqueness part of Lemma 3.1 we have  $\mathcal{E}_1 = D\mathcal{E}$  for some diagonal  $D$  or  $\mathcal{E}T_{\mu} = D\mathcal{E}$ . Considering the diagonal elements we see that  $D = A_{\mu} = \text{diag}(\lambda_{\mu p}, p \in \mathcal{P}_n)$ . Therefore  $\mathcal{E}T_{\mu} = A_{\mu}\mathcal{E}$  for all  $\mu$ . ■

We defined zonal polynomials by defining their coefficients. From a little bit more abstract viewpoint they are eigenfunctions of the linear operator  $\tau_{\nu}$  and the results in this section can be summarized as follows.

**Theorem 3.1.** *Let  $y_p$  be a zonal polynomial then*

$$(3.32) \quad (\tau_{\nu} y_p)(A) = \mathcal{E}_W y_p(AW) = \lambda_{\nu p} y_p(A),$$

where  $W \sim \mathcal{W}(I_k, \nu)$ ,  $A$  is symmetric and  $\lambda_{\nu p}$  is given in (3.21). Conversely (3.32) (for all sufficiently large  $\nu$ ) implies that  $y_p$  is a zonal polynomial.

*Proof:*  $\mathcal{Y} = (\mathcal{Y}_{(n)}, \mathcal{Y}_{(n-1,1)}, \dots, \mathcal{Y}_{(1^n)})' = \mathcal{E}\mathcal{U}$ . Hence by Lemma 3.4

$$\begin{aligned}
 \varepsilon_W\{\mathcal{Y}(AW)\} &= \varepsilon_W\{\mathcal{E}\mathcal{U}(AW)\} \\
 &= \mathcal{E}\varepsilon_W\{\mathcal{U}(AW)\} \\
 &= \mathcal{E}T_\nu\mathcal{U}(A) \\
 &= \Lambda_\nu\mathcal{E}\mathcal{U}(A) \\
 &= \Lambda_\nu\mathcal{Y}(A).
 \end{aligned}
 \tag{3.33}$$

Therefore (3.32) holds. Conversely suppose that (3.32) holds. Let  $\mathcal{Y}_p = \sum_{q \in \mathcal{P}_n} a_q \mathcal{U}_q$ . Then (3.32) implies

$$\mathbf{a}' T_\nu = \lambda_{\nu p} \mathbf{a}',$$

where  $\mathbf{a}' = (a_{(n)}, \dots, a_{(1^n)})$ . Now by the uniqueness part of Lemma 3.4  $\mathbf{a}'$  coincides with the “ $p$ -th” row of  $\mathcal{E}$  up to a multiplicative constant. Therefore  $\mathcal{Y}_p$  is a zonal polynomial.

■

**Corollary 3.3.**

$$(3.34) \quad \varepsilon_W \mathcal{Y}_p(AW) = \lambda_{\nu p} \mathcal{Y}_p(A\Sigma),$$

where  $W$  is distributed according to  $\mathcal{W}(\Sigma, \nu)$ .

*Proof:* let  $W = \Sigma^{\frac{1}{2}} W_1 \Sigma^{\frac{1}{2}}$  then  $W_1$  is distributed according to  $\mathcal{W}(I_k, \nu)$ . Therefore

$$\begin{aligned}
 \varepsilon_W \mathcal{Y}_p(AW) &= \varepsilon_W \mathcal{Y}_p(A\Sigma^{\frac{1}{2}} W_1 \Sigma^{\frac{1}{2}}) \\
 &= \varepsilon_W \mathcal{Y}_p(\Sigma^{\frac{1}{2}} A \Sigma^{\frac{1}{2}} W_1) = \lambda_{\nu p} \mathcal{Y}_p(\Sigma^{\frac{1}{2}} A \Sigma^{\frac{1}{2}}) = \lambda_{\nu p} \mathcal{Y}_p(A\Sigma).
 \end{aligned}
 \tag{3.35}$$

■

The converse part of Theorem 3.1 will be used frequently to show that a particular symmetric polynomial is a zonal polynomial. See Sec. 3.3, Sec. 4.4, and Sec. 4.7.

## §3.2 Integral identities involving zonal polynomials

In addition to (3.32) the zonal polynomials satisfy other integral identities. The fundamental one (Theorem 3.2) is related to the uniform distribution of orthogonal matrices.

The idea of “averaging with respect to the uniform distribution of orthogonal matrices” or “averaging over orthogonal group” was a very important idea of James for the motivation of introducing the zonal polynomials. Therefore it seems worthwhile here to develop the uniform distribution of orthogonal matrices in a constructive way, although a general theory of Haar measures readily establishes it (see Nachbin(1965), Halmos(1974) for example).

**Definition 3.2.** A random orthogonal matrix  $H$  is said to have the *Haar invariant* distribution or the *uniform distribution* if the distribution of  $H\Gamma$  is the same for every orthogonal  $\Gamma$ .

See Anderson(1958), Chapter 13. More formally

**Definition 3.2'.** A probability measure  $P$  on the Borel field of orthogonal matrices is *Haar invariant* if

$$(3.36) \quad P(A) = P(A\Gamma)$$

for every orthogonal  $\Gamma$  and every Borel set  $A$ .

We now examine existence and uniqueness of Haar invariant distribution. We settle the uniqueness question first.

**Lemma 3.6.** Let two probability measures  $P_1, P_2$  satisfy (3.36). Then  $P_1(A) = P_2(A)$  for every Borel set  $A$ . Furthermore  $P_1(A) = P_1(A')$  where  $A' = \{H' \mid H \in A\}$ .

*Proof:* Let  $H_1, H_2$  be independently distributed according to  $P_1, P_2$  respectively. Then

$$(3.37) \quad Pr(H_1 H_2' \in A) = \mathcal{E}_{H_2}\{Pr(H_1 H_2' \in A \mid H_2)\} = \mathcal{E}_{H_2}\{P_1(A)\} = P_1(A).$$

Similarly

$$(3.38) \quad Pr(H_1 H_2' \in A) = Pr(H_2 H_1' \in A') = \mathcal{E}_{H_1}\{Pr(H_2 H_1' \in A' \mid H_1)\} = P_2(A').$$

Hence

$$(3.39) \quad P_1(A) = P_2(A').$$

Putting  $P_1 = P_2$  we obtain  $P_1(A) = P_1(A'), P_2(A) = P_2(A')$ . Substituting this into (3.39) we obtain  $P_1(A) = P_2(A)$ . ■

**Remark 3.6.** For a more rigorous proof (3.37) and (3.38) have to be converted to the form of Fubini's theorem, as is done in standard proofs (see Section 60 of Halmos(1974)). However in the latter form the change of variables and the change of order of integration seems to be harder to understand intuitively. The same remark applies to the proof of Lemma 3.8 below.

**Remark 3.7.** If  $H$  has the uniform distribution then for every orthogonal  $H_1$ ,  $H_1 H$  has the uniform distribution. This follows from the fact that  $H_1 H$  and  $H_1 H \Gamma$  have the same distribution for every orthogonal  $\Gamma$ . The second assertion of Lemma 3.6 shows that if  $H$  is uniform then  $H'$  is uniform.

Existence can be very explicitly established as follows.

**Lemma 3.7.** (Saw(1970)) Let  $U = (u_{ij})$  be a  $k \times k$  matrix such that  $u_{ij}$  are independent standard normal variables. Then with probability 1,  $U$  can be uniquely expressed as

$$(3.40) \quad U = TH,$$

where  $T = (t_{ij})$  is lower triangular with positive diagonal elements and  $H$  is orthogonal. Furthermore (i)  $T$  and  $H$  are independent, (ii)  $H$  is uniform, (iii)  $t_{ij}$  are all independent and  $t_{ii} \sim \chi(k - i + 1)$ ,  $t_{ij} \sim \mathcal{N}(0, 1)$ ,  $i > j$ .

*Proof:*  $U$  is nonsingular with probability 1. Therefore suppose  $|U| \neq 0$ . Now performing the Gram-Schmidt orthogonalization to the rows of  $U$  starting from the first row we obtain  $SU = H$  where  $S$  is lower triangular with positive diagonal elements and  $H$  is orthogonal. Letting  $T = S^{-1}$  we obtain (3.40). Since (3.40) corresponds to the uniquely defined Gram-Schmidt orthogonalization  $T, H$  are unique. Now  $W = UU' = TT'$  is distributed according to  $\mathcal{W}(I_k, k)$ . Hence (iii) follows from Lemma 3.3. To show (i) and (ii) we first note that for any orthogonal  $\Gamma$ ,  $U\Gamma$  has the same distribution as  $U$ . Furthermore  $U\Gamma = T(H\Gamma)$ . Therefore  $H\Gamma$  is the resulting orthogonal matrix obtained by performing Gram-Schmidt orthogonalization to the rows of  $U\Gamma$  and  $T$  is common to  $U$  and  $U\Gamma$ . This implies that given  $T$  the conditional distributions of  $H$  and  $H\Gamma$  are the same. Therefore the conditional distribution of  $H$  given  $T$  is uniform. Now by unconditioning we see that  $T$  and  $H$  are independent and  $H$  has the uniform distribution. ■

Now we prove the following fundamental identity (James (1961)). The proof is only a slight modification of one in Saw(1977).

**Theorem 3.2.** *Let  $A, B$  be  $k \times k$  symmetric matrices. Then*

$$(3.41) \quad \mathcal{E}_H \mathcal{Y}_p(AHBH') = \mathcal{Y}_p(A) \mathcal{Y}_p(B) / \mathcal{Y}_p(I_k),$$

where  $k \times k$  orthogonal  $H$  has the uniform distribution.

*Proof:* Let  $\lambda(A) = \alpha = (\alpha_1, \dots, \alpha_k)$  and  $\lambda(B) = \beta = (\beta_1, \dots, \beta_k)$ . Let  $A = H_1 D_1 H_1'$ ,  $B = H_2 D_2 H_2'$ , where  $H_1, H_2$  are orthogonal and  $D_1 = \text{diag}(\alpha_1, \dots, \alpha_k)$ ,  $D_2 = \text{diag}(\beta_1, \dots, \beta_k)$ . Now

$$(3.42) \quad \begin{aligned} \mathcal{Y}_p(AHBH') &= \mathcal{Y}_p(H_1 D_1 H_1' H H_2 D_2 H_2' H') \\ &= \mathcal{Y}_p(D_1 H_3 D_2 H_3'), \end{aligned}$$

where  $H_3 = H_1' H H_2$  which has the uniform distribution (see Remark 3.7). Therefore

$$(3.43) \quad \mathcal{E}_H \mathcal{Y}_p(AHBH') = \mathcal{E}_H \mathcal{Y}_p(D_1 H D_2 H').$$

This depends only on  $\alpha = (\alpha_1, \dots, \alpha_k)$ , and  $\beta = (\beta_1, \dots, \beta_k)$ . Now for any permutation matrix  $P$ ,  $A = (H_1 P) (P D_1 P) (P H_1')$ . Noting that a permutation matrix is orthogonal we get  $\mathcal{E}_H \mathcal{Y}_p(D_1 H D_2 H') = \mathcal{E}_H \mathcal{Y}_p(P D_1 P H D_2 H')$ . therefore (3.43) is symmetric in  $\alpha_1, \dots, \alpha_k$ . Similarly it is symmetric in  $\beta_1, \dots, \beta_k$ . Furthermore given  $\beta = (\beta_1, \dots, \beta_k)$ ,  $\mathcal{U}_p(D_1 H D_2 H')$  is a homogeneous polynomial of degree  $n$  in  $(\alpha_1, \dots, \alpha_k)$  (see(3.15)). Therefore taking expectation

$$(3.44) \quad \mathcal{E}_H \mathcal{Y}_p(D_1 H D_2 H') \in V_n(A) \quad \text{for each } \beta = (\beta_1, \dots, \beta_k),$$

and we can write

$$(3.45) \quad \mathcal{E}_H \mathcal{Y}_p(D_1 H D_2 H') = \sum_{q \in \mathcal{P}_n} c_q(\beta) \mathcal{Y}_q(\alpha).$$

Now looking at (3.45) as a function of  $\beta$  (for fixed  $\alpha$ ) we see that  $c_q(\beta) \in V_n(B)$ . Hence

$$(3.46) \quad c_q(\beta) = \sum_{q' \in \mathcal{P}_n} c_{qq'} \mathcal{Y}_{q'}(\beta).$$

Substituting (3.46) into (3.45) and returning to symmetric  $A, B$  we obtain

$$(3.47) \quad \mathcal{E}_H \mathcal{Y}_p(AHBH') = \sum_{q, q'} c_{qq'} \mathcal{Y}_q(A) \mathcal{Y}_{q'}(B).$$

Note that  $c_{qq'} = c_{q'q}$  because  $\mathcal{Y}_p(AHBH') = \mathcal{Y}_p(BH'AH)$  and  $H'$  has the uniform distribution. Now let  $A$  be distributed according to  $\mathcal{W}(\Sigma, \nu_0)$  where  $\nu_0$  is such that  $\lambda_{\nu_0 p}$ ,  $p \in \mathcal{P}$  are all different (see the proof of Lemma 3.4). Then by Corollary 3.3

$$(3.48) \quad \mathcal{E}_A \mathcal{E}_H \mathcal{Y}_p(AHBH') = \sum_{q, q'} c_{qq'} \lambda_{\nu_0 q} \mathcal{Y}_q(\Sigma) \mathcal{Y}_{q'}(B).$$

On the other hand taking expectation with respect to  $A$  first we obtain

$$(3.49) \quad \begin{aligned} \mathcal{E}_H \mathcal{E}_A \mathcal{Y}_p(AHBH') &= \lambda_{\nu_0 p} \mathcal{E}_H \mathcal{Y}_p(\Sigma HBH') \\ &= \lambda_{\nu_0 p} \sum_{q, q'} c_{qq'} \mathcal{Y}_q(\Sigma) \mathcal{Y}_{q'}(B). \end{aligned}$$

Therefore

$$(3.50) \quad 0 = \sum_{q, q'} (\lambda_{\nu_0 p} - \lambda_{\nu_0 q}) c_{qq'} \mathcal{Y}_q(\Sigma) \mathcal{Y}_{q'}(B).$$

This holds for any  $\Sigma$  and  $B$ . Hence  $(\lambda_{\nu_0 p} - \lambda_{\nu_0 q}) c_{qq'} = 0$  for all  $q, q'$ . Since  $\lambda_{\nu_0 p} \neq \lambda_{\nu_0 q}$  for  $p \neq q$  we have  $c_{qq'} = 0$  for all  $q' \neq p$  and all  $q \neq p$ . But  $c_{qq'} = c_{q'q}$ . Therefore  $c_{qq'} = 0$  unless  $q = q' = p$ . Therefore

$$(3.51) \quad \mathcal{E}_H \mathcal{Y}_p(AHBH') = c_{pp} \mathcal{Y}_p(A) \mathcal{Y}_p(B).$$

Putting  $B = I_k$  we obtain

$$(3.52) \quad \mathcal{Y}_p(A) = c_{pp} \mathcal{Y}_p(I_k) \mathcal{Y}_p(A).$$

Hence  $c_{pp} \mathcal{Y}_p(I_k) = 1$  and this proves the theorem. ■

For more about this proof see Section 4.1.

Theorem 3.2 implies the following rather strong result.



**Theorem 3.3.** *Suppose that a  $k \times k$  random symmetric matrix  $V$  has such a distribution that for every orthogonal  $\Gamma$ ,  $\Gamma V \Gamma'$  has the same distribution as  $V$ . Then for symmetric  $A$*

$$(3.53) \quad \mathcal{E}_V \mathcal{Y}_p(AV) = c_p \mathcal{Y}_p(A),$$

where

$$(3.54) \quad c_p = \mathcal{E}_V \{\mathcal{Y}_p(V)\} / \mathcal{Y}_p(I_k).$$

*Proof:* As in the proof of Lemma 3.2  $\mathcal{E}_V \mathcal{Y}_p(AV) \in V_n(A)$ . Now since the distribution of  $\Gamma V \Gamma'$  is the same as  $V$  we have

$$(3.55) \quad \mathcal{E}_V \mathcal{Y}_p(A \Gamma V \Gamma') = \mathcal{E}_V \mathcal{Y}_p(AV).$$

Letting  $\Gamma$  be uniformly distributed

$$(3.56) \quad \begin{aligned} \mathcal{E}_V \mathcal{Y}_p(AV) &= \mathcal{E}_\Gamma \mathcal{E}_V \mathcal{Y}_p(A \Gamma V \Gamma') \\ &= \mathcal{E}_V \mathcal{E}_\Gamma \mathcal{Y}_p(A \Gamma V \Gamma') \\ &= \mathcal{E}_V \{\mathcal{Y}_p(A) \mathcal{Y}_p(V) / \mathcal{Y}_p(I_k)\} \\ &= \mathcal{Y}_p(A) \mathcal{E}_V \{\mathcal{Y}_p(V)\} / \mathcal{Y}_p(I_k). \end{aligned}$$

■

**Remark 3.8.** *In the sequel we call the distribution of  $V$  "orthogonally invariant" if it satisfies the condition of Theorem 3.3.*

Although Theorem 3.3 has not been explicitly stated, it has been implicitly used for several cases; first with the multivariate beta distribution by Constantine(1963), later with the inverted Wishart distribution by Khatri(1966) etc. These cases will be examined in Section 4.3 together with the evaluation of  $c_p$  for each case.

We note that Theorem 3.3 is a generalization of Theorem 3.1. If we examine the proof of Theorem 3.2 closely we find that we can take  $A = \Sigma^{\frac{1}{2}} V \Sigma^{\frac{1}{2}}$  in (3.48) and (3.49) where  $V$  has any orthogonally invariant distribution provided that  $c_p, p \in \mathcal{P}_n$ , are all distinct for that distribution. Furthermore the construction of zonal polynomials in Section 3.1 works with any such distribution. Although the Wishart distribution seems to be a natural candidate, we could have used any such distribution from a purely logical point of view.

Orthogonally invariant distributions are characterized as follows.

**Lemma 3.8.** *Let  $V = HDH'$  where  $H$  is orthogonal and  $D$  is diagonal. Let  $H$  and  $D$  be independently distributed such that  $H$  has the uniform distribution. (Diagonal elements of  $D$  can have any distribution.) Then  $V$  has an orthogonally invariant distribution. Conversely all orthogonally invariant distributions can be obtained in this way.*

*Proof:* The first part of the lemma is obvious. To prove the converse suppose that  $V$  has an orthogonally invariant distribution. Now we form a new random matrix  $\tilde{V} = HVH'$  where  $H$  has the uniform distribution independently of  $V$ . Then  $\tilde{V}$  has the same distribution as  $V$  because for any Borel set  $A$

$$\begin{aligned} Pr(\tilde{V} \in A) &= \mathcal{E}_H\{Pr(HVH' \in A | H)\} \\ (3.57) \quad &= \mathcal{E}_H\{Pr(V \in A)\} = Pr(V \in A). \end{aligned}$$

Now we evaluate  $Pr(\tilde{V} \in A)$  by conditioning on  $V$ . For fixed  $V$  we can write  $V = \Gamma D \Gamma'$  where  $\Gamma$  is orthogonal and  $D = \text{diag}(d_1, \dots, d_k)$ . We require  $d_1 \geq \dots \geq d_k$  then  $D = D(V)$  is unique. Then

$$\begin{aligned} Pr(\tilde{V} \in A | V) &= Pr(H\Gamma D(V)\Gamma' H' \in A | V) \\ (3.58) \quad &= Pr(HD(V)H' \in A | V). \end{aligned}$$

Note that we replaced  $H\Gamma$  by  $H$  since  $H\Gamma$  has the uniform distribution. Hence

$$\begin{aligned} Pr(\tilde{V} \in A) &= \mathcal{E}_V\{Pr(\tilde{V} \in A | V)\} \\ (3.59) \quad &= \mathcal{E}_V\{Pr(HD(V)H' \in A | V)\} \\ &= Pr(HD(V)H' \in A). \end{aligned}$$

This proves the lemma. ■

**Remark 3.9.** *Note that the set of orthogonally invariant distributions is convex with respect to taking mixture of distributions. Lemma 3.8 implies that the extreme points of this convex set are given by those distributions for which  $D$  is degenerate.*

We can replace  $H$  in Theorem 3.2 by  $U$  whose elements are independent normal variables.

**Theorem 3.4.** *Let  $U = (u_{ij})$  be a  $k \times k$  matrix such that  $u_{ij}$  are independent standard normal variables. Then for symmetric  $A, B$*

$$(3.60) \quad \varepsilon_U \mathcal{Y}_p(AUBU') = \frac{\lambda_{kp}}{\mathcal{Y}_p(I_k)} \mathcal{Y}_p(A) \mathcal{Y}_p(B).$$

*Proof:* By Lemma 3.7  $U = TH$ . Then

$$(3.61) \quad \begin{aligned} \varepsilon_U \mathcal{Y}_p(AUBU') &= \varepsilon_T \varepsilon_H \mathcal{Y}_p(ATHBH'T') \\ &= \varepsilon_T \varepsilon_H \mathcal{Y}_p(T'ATHBH') \\ &= \varepsilon_T \mathcal{Y}_p(T'AT) \mathcal{Y}_p(B) / \mathcal{Y}_p(I_k) \\ &= \varepsilon_T \mathcal{Y}_p(ATT') \mathcal{Y}_p(B) / \mathcal{Y}_p(I_k) \\ &= \frac{\lambda_{kp}}{\mathcal{Y}_p(I_k)} \mathcal{Y}_p(A) \mathcal{Y}_p(B). \end{aligned}$$

We used the fact that  $TT' = UU' \sim \mathcal{W}(I_k, k)$ . ■

Theorem 3.4 leads to the following important observation.

**Theorem 3.5.**  $b_p = \lambda_{kp} / \mathcal{Y}_p(I_k)$  is a constant independent of  $k$ .

*Proof:* . Let  $A, B$  be augmented by zeros as

$$\tilde{A} = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}, \quad k_1 \times k_1, \quad \tilde{B} = \begin{pmatrix} B & 0 \\ 0 & 0 \end{pmatrix}, \quad k_1 \times k_1.$$

Then  $\mathcal{Y}_p(\tilde{A}) = \mathcal{Y}_p(A)$ ,  $\mathcal{Y}_p(\tilde{B}) = \mathcal{Y}_p(B)$ , and  $\mathcal{Y}_p(\tilde{A}\tilde{U}\tilde{B}\tilde{U}') = \mathcal{Y}_p(AUBU')$  where  $\tilde{U} (k_1 \times k_1)$  is obtained by adding independent standard normal variables to  $U$ . Therefore

$$(3.62) \quad \begin{aligned} \frac{\lambda_{kp}}{\mathcal{Y}_p(I_k)} \mathcal{Y}_p(A) \mathcal{Y}_p(B) &= \varepsilon_U \mathcal{Y}_p(AUBU') \\ &= \varepsilon_{\tilde{U}} \mathcal{Y}_p(\tilde{A}\tilde{U}\tilde{B}\tilde{U}') \\ &= \frac{\lambda_{k_1 p}}{\mathcal{Y}_p(I_{k_1})} \mathcal{Y}_p(\tilde{A}) \mathcal{Y}_p(\tilde{B}) \\ &= \frac{\lambda_{k_1 p}}{\mathcal{Y}_p(I_{k_1})} \mathcal{Y}_p(A) \mathcal{Y}_p(B). \end{aligned}$$

Hence  $\lambda_{kp} / \mathcal{Y}_p(I_k) = \lambda_{k_1 p} / \mathcal{Y}_p(I_{k_1})$ . ■

We evaluate the  $b_p$ 's in Section 4.2. Corresponding to Theorem 3.3, Theorem 3.4 can be generalized as follows.

**Theorem 3.6.** *Let  $X$  be a  $k \times k$  random matrix (not symmetric) such that for every orthogonal  $\Gamma_1, \Gamma_2$ , the distribution of  $\Gamma_1 X \Gamma_2$  is the same as the distribution of  $X$ . Then for symmetric  $A, B$*

$$(3.63) \quad \mathcal{E}_X y_p(AXBX') = \gamma_p y_p(A) y_p(B),$$

where

$$(3.64) \quad \gamma_p = \mathcal{E}_X \{y_p(XX')\} / \{y_p(I_k)\}^2.$$

*Proof:* For any orthogonal  $\Gamma_1$

$$(3.65) \quad \mathcal{E}_X y_p(AXBX') = \mathcal{E}_X y_p(AX\Gamma_1 B\Gamma_1' X').$$

Letting  $\Gamma_1$  be uniformly distributed we obtain

$$(3.66) \quad \begin{aligned} \mathcal{E}_X y_p(AXBX') &= \frac{y_p(B)}{y_p(I_k)} \mathcal{E}_X y_p(X'AX) \\ &= \frac{y_p(B)}{y_p(I_k)} \mathcal{E}_X y_p(AXX'). \end{aligned}$$

Now  $V = XX'$  has an orthogonally invariant distribution because  $\Gamma_2 V \Gamma_2' = (\Gamma_2 X)(\Gamma_2 X)'$ . Therefore by Theorem 3.3

$$(3.67) \quad \mathcal{E}_X y_p(AXX') = y_p(A) \mathcal{E}_X \{y_p(XX')\} / y_p(I_k).$$

Substituting (3.67) into (3.66) we obtain the lemma. ■

**Remark 3.10.** *We call the distribution of  $X$  "orthogonally biinvariant" if it satisfies the condition of Theorem 3.6.*

Corresponding to Lemma 3.8 we have

**Lemma 3.9.** *Let  $X = H_1 D H_2$  where  $H_1, H_2$  are orthogonal and  $D$  is diagonal. Let  $H_1, H_2, D$ , be independently distributed such that  $H_1, H_2$  have the uniform distribution. ( $D$  can have any distribution.) Then  $X$  has an orthogonally biinvariant distribution. Conversely all orthogonally biinvariant distributions can be obtained in this way.*

The proof is entirely analogous to the proof of Lemma 3.8, therefore we omit it.

**Remark 3.11.** *The notion of orthogonal biinvariance can be applied to rectangular matrices. If  $X$  is  $k \times m$  in Theorem 3.6 we obtain*

$$(3.68) \quad \gamma_p = \frac{\mathcal{E}_X \mathcal{Y}_p(XX')}{\mathcal{Y}_p(I_k) \mathcal{Y}_p(I_m)}$$

and in Lemma 3.9 (for  $k \leq m$ ) we replace  $X = H_1 D H_2$  by  $X = H_1(D, 0) H_2$ .

In the sequel we almost exclusively work with the Wishart and the normal distributions. But in view of Theorem 3.3 and Theorem 3.6 there could be other distributions which give information on various aspects of zonal polynomials.

### §3.3 An integral representation of zonal polynomials

We prove an integral representation by Kates(1980) which shows that (i) zonal polynomials are positive for positive definite  $A$  and increasing in each root of  $A$ , (ii) in the normalization  $Z_p$  defined below the coefficients  $a_{pq}$  in  $Z_p = \sum a_{pq} M_q$  are nonnegative integers. The derivation by Kates is rather abstract but the integral representation can be proved directly in our framework. The representation can be formulated in several ways. James(1973) derived one involving uniform orthogonal matrix. We discuss these variations in Section 4.7.

From Theorem 3.5 we see that a constant  $b_p$  or equivalently the value of a zonal polynomial at  $I_k$  describes a particular normalization. The normalization  $Z_p$  is the simplest one in this sense.

**Definition 3.3.** A particular normalization of a zonal polynomial denoted by  $Z_p$  is defined by

$$(3.69) \quad Z_p(I_k) = \lambda_{kp},$$

or  $b_p = 1$  in Theorem 3.5.

**Theorem 3.7.** (Kates, 1980) Let  $p = (p_1, \dots, p_\ell)$ . For  $k \times k$  symmetric  $A$

$$(3.70) \quad Z_p(A) = \mathcal{E}_U \{ \Delta_1^{p_1 - p_2} \Delta_2^{p_2 - p_3} \dots \Delta_\ell^{p_\ell} \},$$

where  $\Delta_i = UAU'(1, \dots, i)$  is the determinant of the upper left minor of  $UAU'$  and  $U$  is a  $k \times k$  random matrix whose entries are independent standard normal variables.

*Proof:* Let

$$(3.71) \quad \tilde{Z}_p(A) = \mathcal{E}_U \{\Delta_1^{p_1 - p_2} \dots \Delta_\ell^{p_\ell}\}.$$

It can be routinely checked that (3.71) is a homogeneous symmetric polynomial of degree  $n = |p|$  in the roots of  $A$ . Furthermore augmenting  $A$  to  $\tilde{A}$  ( $k_1 \times k_1$ ) by adding zeros and augmenting  $U$  to  $\tilde{U}$  by adding independent standard normal variables does not change the upper left part of  $UAU'$ . Therefore (3.71) does not depend on  $k$ . Hence  $\tilde{Z}_p(A) \in V_n(A)$ . Now we want to show

$$(3.72) \quad \tau_\nu \{\tilde{Z}_p(A)\} = \lambda_{\nu p} \tilde{Z}_p(A).$$

Let

$$\tilde{A} = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} : \nu \times \nu$$

and  $\tilde{W} = Y'Y$  where  $Y$  is a  $\nu \times \nu$  matrix whose entries are standard normal variables. Then

$$(3.73) \quad \begin{aligned} \tau_\nu \{\tilde{Z}_p(A)\} &= \mathcal{E}_{\tilde{W}} \{\tilde{Z}_p(\tilde{A}\tilde{W})\} \\ &= \mathcal{E}_Y \mathcal{E}_{\tilde{U}} \left\{ \prod_{i=1}^{\ell} [\tilde{U}Y\tilde{A}Y'\tilde{U}'(1, \dots, i)]^{p_i - p_{i+1}} \right\} \\ &= \mathcal{E}_{\tilde{U}} \mathcal{E}_Y \left\{ \prod_{i=1}^{\ell} [Y\tilde{U}\tilde{A}\tilde{U}'Y'(1, \dots, i)]^{p_i - p_{i+1}} \right\}. \end{aligned}$$

We switched  $\tilde{U}$  and  $Y$  because they have the same distribution. Now by Lemma 3.7  $Y = TH$  and  $H$  can be absorbed into  $U$ . Therefore

$$(3.74) \quad \begin{aligned} &\mathcal{E}_{\tilde{U}} \mathcal{E}_Y \left\{ \prod_{i=1}^{\ell} [Y\tilde{U}\tilde{A}\tilde{U}'Y'(1, \dots, i)]^{p_i - p_{i+1}} \right\} \\ &= \mathcal{E}_{\tilde{U}} \mathcal{E}_T \left\{ \prod_{i=1}^{\ell} [T\tilde{U}\tilde{A}\tilde{U}'T'(1, \dots, i)]^{p_i - p_{i+1}} \right\} \\ &= \mathcal{E}_{\tilde{U}} \mathcal{E}_T \left\{ \prod_{i=1}^{\ell} (t_{11}^2 \dots t_{ii}^2)^{p_i - p_{i+1}} [\tilde{U}\tilde{A}\tilde{U}'(1, \dots, i)]^{p_i - p_{i+1}} \right\} \\ &= \lambda_{\nu p} \mathcal{E}_{\tilde{U}} \left\{ \prod_{i=1}^{\ell} [\tilde{U}\tilde{A}\tilde{U}'(1, \dots, i)]^{p_i - p_{i+1}} \right\} \\ &= \lambda_{\nu p} \tilde{Z}_p(A). \end{aligned}$$

Hence  $\tilde{Z}_p(\mathbf{A})$  is a zonal polynomial by Theorem 3.1. Putting  $\mathbf{A} = \mathbf{I}_k$  we obtain

$$(3.75) \quad \tilde{Z}_p(\mathbf{I}_k) = \mathcal{E}_W \{W(1)^{p_1-p_2} \dots W(1, \dots, \ell)^{p_\ell}\},$$

where  $W \sim \mathcal{W}(\mathbf{I}_k, k)$ . Again by the triangular decomposition  $W = \mathbf{T}\mathbf{T}'$  (Lemma 3.3) we obtain  $\tilde{Z}_p(\mathbf{I}_k) = \lambda_{kp}$ . Therefore  $\tilde{Z}_p(\mathbf{A}) = Z_p(\mathbf{A})$ . ■

Note that the coefficients of the monomial terms in  $Z_p$  are integers, being the expected value of sum of products of independent standard normal variables. Furthermore if  $\mathbf{A} = \text{diag}(\alpha_1, \dots, \alpha_k)$  then by the Binet-Cauchy theorem (see Gantmacher(1959) for example)

$$(3.76) \quad \begin{aligned} \mathbf{U}\mathbf{A}\mathbf{U}'(1, \dots, r) &= \sum_{i_1 < \dots < i_r} \sum_{j_1 < \dots < j_r} U \begin{pmatrix} 1, \dots, r \\ i_1, \dots, i_r \end{pmatrix} \mathbf{A} \begin{pmatrix} i_1, \dots, i_r \\ j_1, \dots, j_r \end{pmatrix} U' \begin{pmatrix} j_1, \dots, j_r \\ 1, \dots, r \end{pmatrix} \\ &= \sum_{i_1 < \dots < i_r} \alpha_{i_1} \dots \alpha_{i_r} \left\{ U \begin{pmatrix} 1, \dots, r \\ i_1, \dots, i_r \end{pmatrix} \right\}^2, \end{aligned}$$

where  $B \begin{pmatrix} i_1, \dots, i_r \\ j_1, \dots, j_r \end{pmatrix}$  denotes the determinant of a minor formed by rows  $i_1, \dots, i_r$  and columns  $j_1, \dots, j_r$  of  $\mathbf{B}$ . (3.76) is obviously increasing in each  $\alpha_i$  when  $\mathbf{A}$  is positive definite. Furthermore coefficients for monomial terms are nonnegative. These points are discussed in Kates(1980). For more about this see Section 4.1. Generalizations of Theorem 3.7 will be discussed in Section 4.7.

### §3.4 A generating function of zonal polynomials

One of the main contributions of Saw(1977) is his generating function which gives a relatively simple way of computing zonal polynomials. Clearly  $(\text{tr } \mathbf{C})^n \in V_n(\mathbf{C})$ . Therefore we can write

$$(3.77) \quad (\text{tr } \mathbf{C})^n = \sum_{p \in \mathcal{P}_n} d_p Z_p(\mathbf{C}).$$

Let  $C = AUBU'$  where  $A = \text{diag}(\alpha_1, \dots, \alpha_k)$ ,  $B = \text{diag}(\beta_1, \dots, \beta_k)$  and the elements of  $U$  are independent standard normal variables. Then by Theorem 3.4

$$\begin{aligned}
 & \mathcal{E}_U(\text{tr } AUBU')^n \\
 &= \sum_{p \in \mathcal{P}_n} d_p \mathcal{E}_U Z_p(AUBU') \\
 (3.78) \quad &= \sum_{p \in \mathcal{P}_n} d_p \frac{\lambda_{kp}}{Z_p(I_k)} Z_p(A) Z_p(B) \\
 &= \sum_{p \in \mathcal{P}_n} d_p Z_p(A) Z_p(B).
 \end{aligned}$$

Therefore for sufficiently small  $\theta$

$$\begin{aligned}
 & \mathcal{E}_U\{\exp(\theta \text{tr } AUBU')\} \\
 &= \mathcal{E}_U\left\{\sum_{n=0}^{\infty} (\theta^n/n!) (\text{tr } AUBU')^n\right\} \\
 (3.79) \quad &= \sum_{n=0}^{\infty} (\theta^n/n!) \sum_{p \in \mathcal{P}_n} d_p Z_p(A) Z_p(B).
 \end{aligned}$$

On the other hand

$$(3.80) \quad \text{tr } AUBU' = \sum_{i,j}^k \alpha_i \beta_j u_{ij}^2.$$

Hence for sufficiently small  $\theta$

$$\begin{aligned}
 (3.81) \quad \mathcal{E}_U\{\exp(\theta \text{tr } AUBU')\} &= \mathcal{E}_U\left\{\exp\left(\theta \sum_{i,j}^k \alpha_i \beta_j u_{ij}^2\right)\right\} \\
 &= \prod_{i,j}^k (1 - 2\theta \alpha_i \beta_j)^{-\frac{1}{2}}.
 \end{aligned}$$

From (3.79) and (3.81) we obtain

**Theorem 3.8.**

$$(3.82) \quad \prod_{i,j}^k (1 - 2\theta \alpha_i \beta_j)^{-\frac{1}{2}} = \sum_{n=0}^{\infty} (\theta^n/n!) \sum_{p \in \mathcal{P}_n} d_p Z_p(A) Z_p(B).$$



The left hand side of (3.82) can be expanded as follows.

$$\begin{aligned}
 & \prod_{i,j}^k (1 - 2\theta\alpha_i\beta_j)^{-\frac{1}{2}} \\
 &= \exp\{\log \prod_{i,j}^k (1 - 2\theta\alpha_i\beta_j)^{-\frac{1}{2}}\} \\
 &= \exp\left\{\frac{1}{2} \sum_{i,j}^k \sum_{r=1}^{\infty} \frac{(2\theta)^r}{r} \alpha_i^r \beta_j^r\right\} \\
 (3.83) \quad &= \exp\left\{\frac{1}{2} \sum_{r=1}^{\infty} \frac{(2\theta)^r}{r} t_r(A) t_r(B)\right\} \\
 &= \sum_{n=0}^{\infty} \sum_{p \in \mathcal{P}_n} (2\theta)^n \tau_p(A) \tau_p(B) \\
 &\quad \cdot \frac{1}{p_1!} \left(\frac{1}{2}\right)^{p_1} \binom{p_1}{p_1 - p_2, p_2 - p_3, \dots, p_{\ell(p)}} \left(\prod_{r=1}^{\ell(p)} r^{p_r - p_{r+1}}\right)^{-1} \\
 &= \sum_{n=0}^{\infty} (\theta^n/n!) \sum_{p \in \mathcal{P}_n} c_p \tau_p(A) \tau_p(B),
 \end{aligned}$$

where

$$(3.84) \quad c_p = |p|! 2^{|p| - h(p)} \left\{ \prod_{r=1}^{\ell(p)} r^{p_r - p_{r+1}} (p_r - p_{r+1})! \right\}^{-1}.$$

The fourth equality follows from the fact that  $\tau_p$  being a product of  $p_1$  terms comes only from the  $p_1$ -th power term in the expansion of exp. Comparing the coefficient of  $\theta^n$  in (3.82) and (3.83) we obtain

$$(3.85) \quad \sum_{p \in \mathcal{P}_n} d_p Z_p(A) Z_p(B) = \sum_{p \in \mathcal{P}_n} c_p \tau_p(A) \tau_p(B).$$

Note that  $c_p$  is positive for every  $p \in \mathcal{P}_n$ . Therefore putting  $A = B$  and regarding (3.85) as a quadratic form in  $V_n(A)$  we see that (3.85) is positive definite. It follows that  $d_p > 0$  for every  $p \in \mathcal{P}_n$ . Now let  $D = \text{diag}(d_p, p \in \mathcal{P}_n)$ ,  $C = \text{diag}(c_p, p \in \mathcal{P}_n)$ . We recall that  $Z = \mathcal{E}\mathcal{U}$  where  $\mathcal{E}$  is upper triangular and  $\mathcal{T} = F\mathcal{U}$  where  $F$  is lower triangular (see(2.37)). Therefore in matrix notation (3.85) is written as

$$(3.86) \quad \mathcal{U}(A)' \mathcal{E}' D \mathcal{E} \mathcal{U}(B) = \mathcal{U}(A)' F' C F \mathcal{U}(B),$$

or

$$(3.87) \quad \mathbf{E}'\mathbf{D}\mathbf{E} = \mathbf{F}'\mathbf{C}\mathbf{F}.$$

We note that the left hand side and the right hand side correspond to two different triangular decompositions of the same symmetric positive definite matrix.  $\mathbf{F}$  can be computed from (2.33) or alternatively  $\mathbf{F}$  can be obtained from tables given in David, Kendall, and Barton (1966) for  $n \leq 12$ . Therefore we can compute the right hand side of (3.85) relatively easily, then we decompose the resulting positive definite matrix as  $\mathbf{E}'\mathbf{D}\mathbf{E}$ . Diagonal elements of  $\mathbf{E}$  corresponding to  $\mathbf{Z}_p$  is obtained in (4.33). This determines  $\mathbf{D}$  and  $\mathbf{E}$  uniquely.

**Remark 3.12.** *Saw (1977) defined zonal polynomials or the coefficient matrix  $\mathbf{E}$  by (3.87) and derived the first part of Theorem 3.1 from this definition. It seems that (3.87) should be looked at as providing a convenient algorithm for obtaining  $\mathbf{E}$  rather than providing a definition of  $\mathbf{E}$  because it lacks the conceptual motivation necessary for a definition.*

Actually  $d_p$  is known to be (James (1964), formula (18))

$$(3.88) \quad \begin{aligned} d_p &= \chi_{[2p]}(1) 2^n n! / (2n)! \\ &= \frac{2^n n! \prod_{i < j} (2p_i - 2p_j - i + j)}{\prod_{i=1}^{\ell(p)} (2p_i + \ell(p) - i)!}, \end{aligned}$$

where  $n = |p|$  and  $\chi_{[2p]}(1) = (2n)! \prod_{i < j} (2p_i - 2p_j - i + j) / \prod_{i=1}^{\ell(p)} (2p_i + \ell(p) - i)!$  is "the dimension of the representation  $(2p) = (2p_1, \dots, 2p_{\ell(p)})$  of the symmetric group on  $2n$  symbols." This is one thing we were unable to obtain by our elementary approach. It was obtained by James (1961) using group representation theory. We will discuss this point again in Section 4.2 and Section 5.4.

$d_p \mathbf{Z}_p$  is usually denoted by  $\mathcal{C}_p$  so that

$$(3.89) \quad (\text{tr } \mathbf{A})^n = \sum_{p \in \mathcal{P}_n} \mathcal{C}_p(\mathbf{A}).$$

This notation often makes it simpler to write down various noncentral densities. Our last item in Chapter 3 is related to this point.

**Lemma 3.10.**

$$(3.90) \quad \begin{aligned} \mathcal{E}_H(\text{tr } \mathbf{A}\mathbf{H})^{2n} &= \sum_{p \in \mathcal{P}_n} \frac{2^n n! d_p}{(2n)! \lambda_{kp}} \mathbf{Z}_p(\mathbf{A}\mathbf{A}') \\ &= \sum_{p \in \mathcal{P}_n} \frac{2^n n!}{(2n)! \lambda_{kp}} \mathcal{C}_p(\mathbf{A}\mathbf{A}'), \end{aligned}$$

where  $k \times k$  orthogonal  $H$  is uniformly distributed.

*Proof:* Let the singular value decomposition be  $A = \Gamma_1 D \Gamma_2$  where  $\Gamma_1, \Gamma_2$  are orthogonal,  $D = \text{diag}(\delta_1, \dots, \delta_k)$  and  $\theta_i = \delta_i^2, i = 1, \dots, k$  are the characteristic roots of  $AA'$ . Then  $(\text{tr } AH)^{2n} = (\text{tr } \Gamma_1 D \Gamma_2 H)^{2n} = (\text{tr } D \Gamma_2 H \Gamma_1)^{2n}$  and  $\Gamma_2 H \Gamma_1$  has the same distribution as  $H$ . Therefore  $\mathcal{E}_H(\text{tr } AH)^{2n}$  is a  $2n$ -th degree homogeneous polynomial in  $\delta_1, \dots, \delta_k$ . Furthermore the order of  $\delta_1, \dots, \delta_k$  and the sign for each  $\delta_i$  are arbitrary in the singular value decomposition. It follows that  $\mathcal{E}_H(\text{tr } AH)^{2n} \in V_n(AA')$ . Therefore we can write

$$(3.91) \quad \mathcal{E}_H(\text{tr } AH)^{2n} = \sum_{p \in \mathcal{P}_n} a_p Z_p(AA').$$

Now let  $A = \text{diag}(\alpha_1, \dots, \alpha_k)$  and  $U = (u_{ij})$  be as before. Then  $\text{tr } AU = \sum \alpha_i u_{ii}$  is distributed according to  $\mathcal{N}(0, \sum \alpha_i^2)$ . Hence

$$(3.92) \quad \begin{aligned} \mathcal{E}_U(\text{tr } AU)^{2n} &= \left( \sum \alpha_i^2 \right)^n \cdot 1 \cdot 3 \cdots (2n-1) \\ &= \frac{(2n)!}{2^n n!} (\text{tr } AA')^n \\ &= \sum_{p \in \mathcal{P}_n} \frac{(2n)! d_p}{2^n n!} Z_p(AA'). \end{aligned}$$

On the other hand by Lemma 3.7

$$(3.93) \quad \begin{aligned} \mathcal{E}_U(\text{tr } AU)^{2n} &= \mathcal{E}_{T,H}(\text{tr } ATH)^{2n} \\ &= \sum_{p \in \mathcal{P}_n} a_p \mathcal{E}_T Z_p(ATT' A') \\ &= \sum_{p \in \mathcal{P}_n} a_p \lambda_{kp} Z_p(AA'). \end{aligned}$$

Comparing (3.92) and (3.93) we obtain (3.90). ■

# More properties of zonal polynomials

This chapter is a collection of results which are for the most part generalizations and refinements of the basic results given in Chapter 3. We do not attempt to collect various known identities involving zonal polynomials. For this purpose the reader is referred to an excellent survey paper by Subrahmaniam(1976). Actually in the discussion of the orthogonally invariant distributions we saw that zonal polynomials satisfy an infinite number of identities. It is a rather frustrating fact that although many identities for zonal polynomials are already known, explicit forms of zonal polynomials are not known.

## §4.1 Majorization ordering

The proof of Theorem 3.2 which played an essential role for the subsequent development in Chapter 3 is not complete as it is. In (3.49) we argued that

$$(4.1) \quad 0 = \sum_{q, q'} (\lambda_{\nu_0 p} - \lambda_{\nu_0 q}) c_{qq'} y_q(\Sigma) y_{q'}(B),$$

for any  $\Sigma, B$  implies  $(\lambda_{\nu_0 p} - \lambda_{\nu_0 q}) c_{qq'} = 0$ . One objection may be that  $\Sigma$  is restricted to be positive semidefinite. But this causes no trouble since (4.1) is a polynomial and a polynomial which is identically zero for nonnegative arguments has to be zero everywhere. A more serious question is that the dimensionality of  $\Sigma$  and  $B$  is fixed to be  $k \times k$  which is the dimensionality of the uniform orthogonal matrix  $H$ . The same objection applies to the proof of Lemma 3.10. (In the case of  $U$  whose entries are independent standard normal

variables we could augment  $U : k \times k$  to  $\tilde{U} : k_1 \times k_1$  without loss of generality and we could avoid this difficulty.) What we have to consider is  $V_{n,k}(x_1, \dots, x_k)$ , the space of  $n$ -th degree homogeneous symmetric polynomials in  $x_1, \dots, x_k$  ( $k$  fixed) and we have to show that  $\{y_p, p \in \mathcal{P}_n\}$  is indeed linearly independent and forms a basis for  $V_{n,k}(x_1, \dots, x_k)$ . To do this we now study homogeneous symmetric polynomials again from the viewpoint of majorization ordering. The following is a refinement of Lemma 2.5.

**Lemma 4.1.**

$$(4.2) \quad u_p = M_p + \sum_{q \prec p, q \neq p} a_{pq} M_q,$$

$$(4.3) \quad M_p = u_p + \sum_{q \prec p, q \neq p} a^{pq} u_q.$$

*Proof:* Let  $\alpha = x_1 = \dots = x_r$ . Then the degree of  $\alpha$  in  $M_q$  is  $q_1 + \dots + q_r$ . The degree of  $\alpha$  in  $u_p$  is  $p_1 + \dots + p_r$  because

$$(4.4) \quad \begin{aligned} & u_p(\alpha, \dots, \alpha, x_{r+1}, \dots) \\ &= (\alpha + \dots)^{p_1 - p_2} (\alpha^2 + \dots)^{p_2 - p_3} \dots \\ & \quad \cdot (\alpha^r + \dots)^{p_r - p_{r+1}} (\alpha^r x_{r+1} \dots)^{p_{r+1} - p_{r+2}} \dots (\alpha^r x_{r+1} \dots x_\ell + \dots)^{p_\ell} \\ &= c \alpha^{(p_1 - p_2) + 2(p_2 - p_3) + \dots + r(p_r - p_{r+1}) + \dots + r p_\ell} + \dots \\ &= c \alpha^{p_1 + \dots + p_r} + \dots \end{aligned}$$

where  $c$  is a term not containing  $\alpha$ . Let  $Q_r = \{q \mid q \prec p \text{ and } q_1 + \dots + q_r > p_1 + \dots + p_r\}$ . Now since the degree of  $\alpha$  in (4.2) is  $p_1 + \dots + p_r$  we have

$$(4.5) \quad \sum_{q \in Q_r} a_{pq} M_q(\alpha, \dots, \alpha, x_{r+1}, \dots) = 0.$$

Now  $M_q(\alpha, \dots, \alpha, x_{r+1}, \dots, x_k)$  are linearly independent if  $k$  is sufficiently large. Therefore  $a_{pq} = 0$  for  $q \in Q_r$ . Repeating this argument for  $r = 1, 2, \dots$  we have

$$(4.6) \quad a_{pq} = 0 \quad \text{if } q \in Q_1 \cup Q_2 \cup \dots.$$

But if  $q$  is not majorized by  $p$  there exists an  $r$  such that  $q \in Q_r$ . Therefore  $a_{pq} = 0$  for every  $q$  which is not majorized by  $p$ . This proves (4.2). (4.3) can be proved similarly. ■

Let  $V_{n,k}(x_1, \dots, x_k)$  be the vector space of  $n$ -th degree homogeneous symmetric polynomials in  $x_1, \dots, x_k$  ( $k$  fixed).

**Lemma 4.2.**  $\{M_p, p \in \mathcal{P}_n, \ell(p) \leq k\}, \{U_p, p \in \mathcal{P}_n, \ell(p) \leq k\}$  are bases of  $V_{n,k}(x_1, \dots, x_k)$ .

*Proof:* Note that  $M_p(x_1, \dots, x_k) = 0$ ,  $U_p(x_1, \dots, x_k) = 0$  for  $p$  such that  $\ell(p) > k$ . Let  $f \in V_{n,k}(x_1, \dots, x_k)$ . Then from (2.11)

$$\begin{aligned} f(x_1, \dots, x_k) &= \sum_{p \in \mathcal{P}_n} a_p M_p(x_1, \dots, x_k) \\ (4.7) \qquad \qquad &= \sum_{p \in \mathcal{P}_n, \ell(p) \leq k} a_p M_p(x_1, \dots, x_k). \end{aligned}$$

Therefore any  $f \in V_{n,k}(x_1, \dots, x_k)$  can be written as a linear combination of  $M_p$ 's for which  $p \in \mathcal{P}_n, \ell(p) \leq k$ . Now suppose

$$(4.8) \qquad \sum_{q \in \mathcal{P}_n, \ell(q) \leq k} a_q M_q(x_1, \dots, x_k) = 0.$$

Then differentiating (4.8) with respect to  $x_i$   $p_i$  times,  $i = 1, \dots, \ell(p)$ , (note  $\ell(p) \leq k$ ) and setting  $0 = x_1 = \dots = x_k$  we obtain  $a_p = 0$ . Therefore  $\{M_p, p \in \mathcal{P}_n, \ell(p) \leq k\}$  is linearly independent in  $V_{n,k}(x_1, \dots, x_k)$ . This shows that  $\{M_p, p \in \mathcal{P}_n, \ell(p) \leq k\}$  is a basis of  $V_{n,k}(x_1, \dots, x_k)$ . To show that  $\{U_p, p \in \mathcal{P}_n, \ell(p) \leq k\}$  is a basis it suffices to observe

$$\begin{aligned} M_p(x_1, \dots, x_k) &= U_p(x_1, \dots, x_k) + \sum_{q < p} a^{pq} U_q(x_1, \dots, x_k) \\ (4.9) \qquad \qquad &= U_p(x_1, \dots, x_k) + \sum_{q < p, \ell(q) \leq k} a^{pq} U_q(x_1, \dots, x_k). \end{aligned}$$

This and (4.7) with  $f$  replaced by  $U_p$  shows that  $\{U_p, p \in \mathcal{P}_n, \ell(p) \leq k\}$  is another basis of  $V_{n,k}(x_1, \dots, x_k)$ . ■

**Remark 4.1.** It is known that  $a_{pq}$  in (4.2) is nonzero and positive if and only if  $q \prec p$ . This is called the Gale-Ryser theorem. (See Macdonald(1979), Marshall and Olkin(1979).)

Now we prove the following.

**Theorem 4.1.**

$$(4.10) \quad y_p = \sum_{q \prec p} a_{pq} u_q = \sum_{q \prec p} a'_{pq} M_q$$

and  $\{y_p, p \in \mathcal{P}_n, \ell(p) \leq k\}$  forms a basis of  $V_{n,k}(x_1, \dots, x_k)$ .

*Proof:* We first note that majorization is transitive, i.e. if  $p^1 \succ p^2, p^2 \succ p^3$  then  $p^1 \succ p^3$ . Therefore in view of Lemma 4.1 the equalities involving  $u$ 's and  $M$ 's are equivalent. Hence we prove one involving  $u$ 's. Now as in the proof of (4.2) the right hand side of (3.16)

$$(4.11) \quad \mathcal{E}_W \left( \sum_{i_1} \alpha_{i_1} W(i_1) \right)^{p_1 - p_2} \left( \sum_{i_1 < i_2} \alpha_{i_1} \alpha_{i_2} W(i_1, i_2) \right)^{p_2 - p_3} \dots$$

has only those monomial terms  $M_q(A)$  for which  $q \prec p$ . Therefore

$$(4.12) \quad \tau_\nu(u_p) = \sum_{q \prec p} b_{pq} M_q.$$

Substituting (4.3) into (4.12) and using the transitivity of majorization we obtain

$$(4.13) \quad \tau_\nu(u_p) = \sum_{q \prec p} b'_{pq} u_q.$$

Now let

$$(4.14) \quad y_p = \sum_{q \leq p} a_{pq} u_q.$$

We want to show that  $Q_p = \{q \mid a_{pq} \neq 0, q \text{ not majorized by } p\}$  is empty. We argue by contradiction. Suppose that  $Q_p$  is nonempty. Let  $q^*$  be the highest partition in  $Q_p$ . Then  $a_{pq} \neq 0$  and  $q > q^*$  imply  $q \prec p$ . For any such  $q$

$$(4.15) \quad a_{pq} \tau_\nu(u_q) = a_{pq} \left\{ \sum_{q' \prec q} b'_{qq'} u_{q'} \right\}.$$

Now  $q' \prec q, q \prec p$  imply  $q' \prec p$ . Hence the right hand side does not have  $u_{q^*}$  term. It follows that  $u_{q^*}$  does not appear in

$$(4.16) \quad \sum_{q^* \prec q \leq p} a_{pq} \tau_\nu(u_q).$$

Obviously

$$(4.17) \quad \sum_{q < q^*} a_{pq} \tau_\nu(u_q)$$

does not involve  $u_{q^*}$  term either. Therefore the coefficient of  $u_{q^*}$  in

$$(4.18) \quad \begin{aligned} \tau_\nu(y_p) &= \sum_{q \leq p} a_{pq} \tau_\nu(u_q) \\ &= a_{pq^*} \tau_\nu(u_{q^*}) + (4.16) + (4.17) \end{aligned}$$

is  $a_{pq^*} \lambda_{\nu q^*}$ . On the other hand

$$(4.19) \quad \tau_\nu(y_p) = \lambda_{\nu p} y_p.$$

Therefore the coefficient of  $u_{q^*}$  on the right hand side of (4.19) is  $\lambda_{\nu p} a_{pq^*}$ . Taking  $\nu = \nu_0$  we have a contradiction. Therefore  $Q_p$  is empty. This proves (4.10).

To prove the second assertion we note that  $q < p$  implies  $\ell(q) \geq \ell(p)$ . Otherwise  $p_1 + \dots + p_{\ell(q)} < p_1 + \dots + p_{\ell(p)} = n = q_1 + \dots + q_{\ell(q)}$  and this contradicts  $q < p$ . Therefore in (4.10) we have only those  $u_q$ 's for which  $\ell(q) \geq \ell(p)$ . Now suppose that  $A$  is  $k \times k$  and  $k < \ell(p)$ . Then every  $u_q(A)$  in (4.10) vanishes. Hence

$$(4.20) \quad y_p(A) = 0 \quad \text{if } A \text{ is } k \times k \text{ and } k < \ell(p).$$

Now write

$$(4.21) \quad u_p = \sum_{q \leq p} a^{pq} y_q.$$

Then

$$(4.22) \quad \begin{aligned} u_p(x_1, \dots, x_k) &= \sum_{q \leq p} a^{pq} y_q(x_1, \dots, x_k) \\ &= \sum_{q \leq p, \ell(q) \leq k} a^{pq} y_q(x_1, \dots, x_k). \end{aligned}$$

Similarly

$$(4.23) \quad y_p(x_1, \dots, x_k) = \sum_{q \leq p, \ell(q) \leq k} a_{pq} u_q(x_1, \dots, x_k).$$

In view of Lemma 4.2, (4.22) and (4.23) imply that  $\{y_p, p \in P_n, \ell(p) \leq k\}$  forms a basis of  $V_{n,k}(x_1, \dots, x_k)$ . ■



In the proofs of Theorem 3.2 and Lemma 3.10 we replace all the summations by

$$(4.24) \quad \sum_{p \in \mathcal{P}_n, \ell(p) \leq k}, \quad \sum_{\substack{q, q' \in \mathcal{P}_n \\ \ell(q) \leq k, \ell(q') \leq k}} \quad \text{etc.}$$

Then those proofs are complete. We do not repeat the steps of those proofs. But in later proofs we will be careful.

**Remark 4.2.** *Using the Gale-Ryser theorem (Remark 4.1) and (3.70) it can be shown that  $a'_{pq}$  in (4.10) is positive if and only if  $q \prec p$ . This is stronger than Theorem 4.1.*

For future references we record (4.20) as a corollary.

**Corollary 4.1.** *If  $A$  is  $k \times k$  and  $\ell(p) > k$  then  $y_p(A) = 0$ .*

#### §4.2 Evaluation of ${}_1y_p(I_k)$

In the sequel we often work with a normalization denoted by  ${}_1y_p$  which has the leading coefficient 1, namely

$$(4.25) \quad {}_1y_p = u_p + \sum_{q \prec p} {}_1a_{pq} u_q.$$

Advantages of this normalization will become clear soon. We shall evaluate  ${}_1y_p(I_k)$ . From Theorem 3.5 we know that  ${}_1b_p = \lambda_{kp} / {}_1y_p(I_k)$  is a constant independent of  $k$ . Therefore our goal is to obtain  ${}_1b_p$ . Now

$$(4.26) \quad {}_1b_p {}_1y_p(I_k) = \lambda_{kp} = Z_p(I_k).$$

Therefore  ${}_1b_p$  is the leading coefficient of  $Z_p$ . This was needed for the unique decomposition of the left hand side of (3.87). We use the following recursive relation.

**Theorem 4.2.** *If  $A$  is a  $k \times k$  symmetric matrix, then*

$$(4.27) \quad |A| {}_1y_p(A) = {}_1y_{p+(1^k)}(A),$$

where  $p + (1^k) = (p_1 + 1, p_2 + 1, \dots, p_k + 1, p_{k+1}, \dots) \in \mathcal{P}_{n+k}$ ,  $n = |p|$ .

*Proof:* If  $\ell(p) > k$  then  ${}_1y_p(A) = 0$  by Corollary 4.1. In this case  $\ell(p + (1^k)) = \ell(p) > k$ . Hence  ${}_1y_{p+(1^k)}(A) = 0$ . (4.27) holds trivially in this case. Now let  $\ell(p) \leq k$ . Let

$(\alpha_1, \dots, \alpha_k) = \lambda(A)$ . Clearly  $|A| {}_1\mathcal{Y}_p(A) \in V_{n+k,k}(\alpha_1, \dots, \alpha_k)$ . Therefore by Theorem 4.1 we can write

$$(4.28) \quad |A| {}_1\mathcal{Y}_p(A) = \sum_{q \in \mathcal{P}_{n+k}, \ell(q) \leq k} a_q {}_1\mathcal{Y}_q(A).$$

Putting  $AW$  in (4.28) and taking expectation with respect to  $\mathcal{W}(I_k, \nu)$  we obtain

$$(4.29) \quad \mathcal{E}_W\{|AW| {}_1\mathcal{Y}_p(AW)\} = \sum_{q \in \mathcal{P}_{n+k}, \ell(q) \leq k} a_q \lambda_{\nu q} {}_1\mathcal{Y}_q(A).$$

Now the left hand side of (4.29) is equal to  $|A| \mathcal{E}_W\{{}_1\mathcal{Y}_p(AW) | W\}$ . Absorbing  $|W|$  into the Wishart density and letting  $\tilde{W}$  be distributed according to  $\mathcal{W}(I_k, \nu + 2)$  we obtain

$$(4.30) \quad \begin{aligned} & |A| \mathcal{E}_W\{{}_1\mathcal{Y}_p(AW) | W\} \\ &= |A| 2^k \prod_{i=1}^k \frac{\Gamma[\frac{1}{2}(\nu + 3 - i)]}{\Gamma[\frac{1}{2}(\nu + 1 - i)]} \mathcal{E}_{\tilde{W}}\{{}_1\mathcal{Y}_p(A\tilde{W})\} \\ &= |A| 2^k \prod_{i=1}^k \frac{\Gamma[\frac{1}{2}(\nu + 3 - i)]}{\Gamma[\frac{1}{2}(\nu + 1 - i)]} \cdot 2^n \prod_{i=1}^{\ell(p)} \frac{\Gamma[p_i + \frac{1}{2}(\nu + 3 - i)]}{\Gamma[\frac{1}{2}(\nu + 3 - i)]} \cdot {}_1\mathcal{Y}_p(A) \\ &= 2^{k+n} \prod_{i=1}^k \frac{\Gamma[p_i + 1 + \frac{1}{2}(\nu + 1 - i)]}{\Gamma[\frac{1}{2}(\nu + 1 - i)]} |A| {}_1\mathcal{Y}_p(A) \\ &= \lambda_{\nu, p+(1^k)} |A| {}_1\mathcal{Y}_p(A) \\ &= \lambda_{\nu, p+(1^k)} \sum_{q \in \mathcal{P}_{n+k}, \ell(q) \leq k} a_q {}_1\mathcal{Y}_q(A). \end{aligned}$$

Subtracting (4.30) from (4.29) we obtain

$$(4.31) \quad \sum_{q \in \mathcal{P}_{n+k}, \ell(q) \leq k} a_q (\lambda_{\nu q} - \lambda_{\nu, p+(1^k)}) {}_1\mathcal{Y}_q(A) = 0.$$

This holds for any  $k \times k$  symmetric  $A$ . Therefore by Theorem 4.1  $a_q (\lambda_{\nu q} - \lambda_{\nu, p+(1^k)}) = 0$  for every  $q$  such that  $q \in \mathcal{P}_{n+k}, \ell(q) \leq k$ . Taking  $\nu = \nu_0$  we obtain  $a_q = 0$  for  $q \neq p + (1^k)$ . Now comparing the leading coefficient in (4.28) we see  $a_{p+(1^k)} = 1$ . This completes the proof. ■

**Corollary 4.2.** (Formula(129) in James(1964)) Let  $p = (p_1, \dots, p_\ell)$  and  $p - (p_\ell^t) = (p_1 - p_\ell, p_2 - p_\ell, \dots, p_{\ell-1} - p_\ell)$ . Then for an  $\ell \times \ell$  symmetric  $A$

$$(4.32) \quad {}_1\mathcal{Y}_p(A) = |A|^{p_\ell} {}_1\mathcal{Y}_{p-(p_\ell^t)}(A).$$

*Proof:*  $|A|^{p_\ell} {}_1\mathcal{Y}_{p-(p_\ell^t)}(A) = |A|^{p_\ell-1} {}_1\mathcal{Y}_{p-(p_\ell^t)+(1^\ell)}(A) = \dots = {}_1\mathcal{Y}_p(A)$ . ■

Applying Corollary 4.2 to the identity matrices of appropriate dimensionalities we can evaluate  ${}_1b_p$  in (4.26).

**Theorem 4.3.**

$$(4.33) \quad {}_1b_p = 2^{|p|} \prod_{i=1}^{\ell(p)} \prod_{j=1}^i \left( \frac{1}{2}i - \frac{1}{2}(j-1) + p_j - p_i \right)_{p_i - p_{i+1}},$$

where  $(a)_k = a(a+1)\cdots(a+k-1)$ .

*Proof:* We prove this by induction on the length of  $p$ . Let  $\ell(p) = 1$ , namely  $p = (p_1)$ . Then

$$(4.34) \quad {}_1y_p(I_1) = {}_1u_p(I_1) = 1^{p_1} = 1.$$

Therefore

$$(4.35) \quad \begin{aligned} {}_1b_p &= \lambda_{1p} / {}_1y_p(I_1) \\ &= 1 \cdot 3 \cdots (2p_1 - 1) \\ &= 2^{p_1} \left( \frac{1}{2} \right)_{p_1}, \end{aligned}$$

which is of the form (4.33). Now suppose that (4.33) is true for  $\ell(p) = k-1$ . We want to show that then (4.33) holds for  $\ell(p) = k$ . Let  $p = (p_1, \dots, p_k)$  and  $p - (p_k^k) = (p_1 - p_k, p_2 - p_k, \dots, p_{k-1} - p_k)$ . Note that  $\ell(p - (p_k^k)) = k-1$ . Putting  $I_k$  in (4.27) we obtain

$$(4.36) \quad {}_1y_p(I_k) = {}_1y_{p-(p_k^k)}(I_k)$$

or

$$(4.37) \quad {}_1b_p = {}_1b_{p-(p_k^k)} \frac{\lambda_{kp}}{\lambda_{k,p-(p_k^k)}}.$$

Using the induction hypothesis

(4.38)

$$\begin{aligned}
{}_1b_p &= 2^{|p-(p_k^k)|} \prod_{i=1}^{k-1} \prod_{j=1}^i \left( \frac{1}{2}i - \frac{1}{2}(j-1) + (p_j - p_k) - (p_i - p_k) \right)_{(p_i - p_k) - (p_{i+1} - p_k)} \cdot \frac{\lambda_{kp}}{\lambda_{k, p-(p_k^k)}} \\
&= 2^{|p-(p_k^k)|} \prod_{i=1}^{k-1} \prod_{j=1}^i \left( \frac{1}{2}i - \frac{1}{2}(j-1) + p_j - p_i \right)_{p_i - p_{i+1}} \\
&\quad \cdot \frac{2^{|p|} \prod_{j=1}^k \Gamma[p_j + \frac{1}{2}(k+1-j)] / \Gamma[\frac{1}{2}(k+1-j)]}{2^{|p-(p_k^k)|} \prod_{j=1}^{k-1} \Gamma[p_j - p_k + \frac{1}{2}(k+1-j)] / \Gamma[\frac{1}{2}(k+1-j)]} \\
&= 2^{|p|} \prod_{i=1}^{k-1} \prod_{j=1}^i \left( \frac{1}{2}i - \frac{1}{2}(j-1) + p_j - p_i \right)_{p_i - p_{i+1}} \cdot \prod_{j=1}^k \left( \frac{1}{2}k - \frac{1}{2}(j-1) + p_j - p_k \right)_{p_k} \\
&= 2^{|p|} \prod_{i=1}^k \prod_{j=1}^i \left( \frac{1}{2}i - \frac{1}{2}(j-1) + p_j - p_i \right)_{p_i - p_{i+1}}.
\end{aligned}$$

Therefore (4.33) holds for  $k = \ell(p)$  and the theorem is proved. ■

There is a curious fact about  ${}_1b_p$ . Let  $k!!$  denote  $1 \cdot 3 \cdots k$  or  $2 \cdot 4 \cdots k$  depending on whether  $k$  is odd or even. Then as above it can be shown by induction that

$$(4.39) \quad {}_1b_p = \prod_{i < j} \frac{(2p_i - 2p_j - i + j - 2)!!}{(2p_i - 2p_j - i + j - 1)!!} \prod_{i=1}^{\ell(p)} (2p_i - i + \ell(p) - 1)!!.$$

Now  $(\text{tr } A)^n = \sum d_p Z_p(A) = \sum d_p {}_1b_p {}_1\mathcal{Y}_p(A)$ . From (3.88)

$$\begin{aligned}
d_p {}_1b_p &= 2^n n! \prod_{i < j} (2p_i - 2p_j - i + j) / \prod_{i=1}^{\ell(p)} (2p_i - i + \ell(p))! \\
(4.40) \quad &\cdot \prod_{i < j} \frac{(2p_i - 2p_j - i + j - 2)!!}{(2p_i - 2p_j - i + j - 1)!!} \prod_{i=1}^{\ell(p)} (2p_i - i + \ell(p) - 1)!! \\
&= 2^n n! \prod_{i < j} \frac{(2p_i - 2p_j - i + j)!!}{(2p_i - 2p_j - i + j - 1)!!} \bigg/ \prod_{i=1}^{\ell(p)} (2p_i - i + \ell(p))!.
\end{aligned}$$

This is very similar to  ${}_1b_p^{-1}$  if we ignore the constant  $2^n n!$ . In Section 5.3 we will see that in the complex case the corresponding quantities  $\tilde{d}_p$ ,  ${}_1\tilde{b}_p$  satisfy an exact relation  $\tilde{d}_p ({}_1\tilde{b}_p)^2 = n!$ .

### §4.3 More on integral identities

In this section we evaluate the constant  $c_p$  in Theorem 3.3 for several distributions. The first one is the inverted Wishart distribution. See Khatri(1966), Constantine(1963).

**Lemma 4.3.** *Let  $W$  be distributed according to  $\mathcal{W}(I_k, \nu)$ ,  $\nu > 2h(p) + k - 1$ . Then*

$$(4.41) \quad \mathcal{E}_W y_p(AW^{-1}) = c_p y_p(A),$$

where

$$(4.42) \quad c_p = \prod_{i=1}^{\ell(p)} \frac{\Gamma[\frac{1}{2}(\nu - k + i) - p_i]}{\Gamma[\frac{1}{2}(\nu - k + i)]2^{p_i}}.$$

*Proof:* Let  $A = \text{diag}(\alpha_1, \dots, \alpha_k)$  without loss of generality. We look at the monomial term  $\alpha^{p_1} \dots \alpha^{p_\ell}$  ( $\ell = \ell(p)$ ). Then as in (3.18) its coefficient in (4.41) is

$$(4.43) \quad \mathcal{E}_W \{W^{-1}(1)^{p_1-p_2} W^{-1}(1, 2)^{p_2-p_3} \dots W^{-1}(1, \dots, \ell)^{p_\ell}\},$$

which has to be equal to  $c_p$ . Let  $W = T'T$  where  $T$  is lower triangular with positive diagonal elements. Then analogous to Lemma 3.3  $t_{ii}, i = 1, \dots, k$ , are independently distributed according to  $\chi(\nu - k + i)$ . Then  $W^{-1} = T^{-1}T'^{-1}$  and  $T^{-1}$  is lower triangular with diagonal elements reciprocal to the diagonal elements of  $T$ . Therefore  $W^{-1}(1, \dots, r) = (t_{11} \dots t_{rr})^{-2}$ .

Hence

$$\begin{aligned} c_p &= \mathcal{E} \{t_{11}^{-2p_1} \dots t_{\ell\ell}^{-2p_\ell}\} \\ &= \prod_{i=1}^{\ell} \frac{\Gamma[\frac{1}{2}(\nu - k + i) - p_i]}{\Gamma[\frac{1}{2}(\nu - k + i)]2^{p_i}} \\ &= \left\{ \prod_{i=1}^{\ell} (\nu - k + i - 2p_i)(\nu - k + i - 2p_i + 2) \dots (\nu - k + i - 2) \right\}^{-1}. \end{aligned}$$

■

Lemma 4.3 implies the following interesting identity which is briefly mentioned in Constantine(1966). Let  $p_{s,t}^*$  be defined by (2.5).

**Lemma 4.4.** *Let  $A$  be an  $t \times t$  positive definite matrix. Then*

$$(4.44) \quad |A|^s \frac{y_p(A^{-1})}{y_p(I_t)} = \frac{y_{p_{s,t}^*}(A)}{y_{p_{s,t}^*}(I_t)},$$

where  $s \geq h(p), t \geq \ell(p)$ .

*Proof:* Without loss of generality let  $A = \text{diag}(\alpha_1, \dots, \alpha_t)$ . Consider  $|A|^s y_p(A^{-1})$ . Let

$$(4.45) \quad y_p(A^{-1}) = \sum_{q \leq p} a_{pq} M_q(1/\alpha_1, \dots, 1/\alpha_t).$$

Note that  $q \leq p$  implies  $h(q) \leq h(p)$ . Now the degree of  $1/\alpha_i$  in  $M_q(1/\alpha_1, \dots, 1/\alpha_t)$  is  $h(q)$ . Hence the degree of  $1/\alpha_i$  in  $y_p(A^{-1})$  is  $h(p)$ . Now  $|A|^s = (\alpha_1 \cdots \alpha_t)^s$  and  $s \geq h(p)$ . We see that  $1/\alpha_i$  is cancelled by  $|A|^s$  and  $|A|^s y_p(A^{-1})$  is a polynomial in  $(\alpha_1, \dots, \alpha_t)$ . Clearly it is symmetric and homogeneous of degree  $st - |p|$ . Therefore  $|A|^s y_p(A^{-1}) \in V_{st-|p|, t}(\alpha_1, \dots, \alpha_t)$ . By Theorem 4.1 we can write

$$(4.46) \quad |A|^s y_p(A^{-1}) = \sum_{q \in \mathcal{P}_{st-|p|}, \ell(q) \leq t} b_q y_q(A).$$

Replacing  $A$  by  $AW$  in (4.46) and taking expectation with respect to  $\mathcal{W}(I_t, \nu)$  we obtain

$$(4.47) \quad \mathcal{E}_W\{|AW|^s y_p(A^{-1}W^{-1})\} = \sum_{q \in \mathcal{P}_{st-|p|}, \ell(q) \leq t} b_q \lambda_{\nu q} y_q(A).$$

Now proceeding as in (4.30) and using Lemma 4.3 the left hand side of (4.47) can be evaluated as

$$(4.48) \quad \begin{aligned} & |A|^s \mathcal{E}_W\{|W|^s y_p(A^{-1}W^{-1})\} \\ &= |A|^s 2^{st} \prod_{i=1}^t \frac{\Gamma[\frac{1}{2}(\nu + 1 + 2s - i)]}{\Gamma[\frac{1}{2}(\nu + 1 - i)]} \mathcal{E}_{\tilde{W}}\{y_p(A^{-1}\tilde{W}^{-1})\} \\ &= |A|^s 2^{st} \prod_{i=1}^t \frac{\Gamma[\frac{1}{2}(\nu + 1 + 2s - i)]}{\Gamma[\frac{1}{2}(\nu + 1 - i)]} \prod_{j=1}^{\ell(p)} \frac{\Gamma[\frac{1}{2}(\nu + 2s - t + j) - p_j]}{\Gamma[\frac{1}{2}(\nu + 2s - t + j)]^{2p_j}} y_p(A^{-1}) \\ &= |A|^s y_p(A^{-1}) 2^{st-|p|} \prod_{i=1}^t \frac{\Gamma[\frac{1}{2}(\nu + 1 + 2s - i)]}{\Gamma[\frac{1}{2}(\nu + 1 - i)]} \\ &\quad \cdot \prod_{j'=1}^t \frac{\Gamma[\frac{1}{2}(\nu + 1 - j') + s - p_{t-j'+1}]}{\Gamma[\frac{1}{2}(\nu + 1 + 2s - j')]} \\ &= |A|^s y_p(A^{-1}) 2^{st-|p|} \prod_{i=1}^t \frac{\Gamma[\frac{1}{2}(\nu + 1 - i) + s - p_{t-i+1}]}{\Gamma[\frac{1}{2}(\nu + 1 - i)]}, \end{aligned}$$

where  $\tilde{W} \sim \mathcal{W}(I_t, \nu + 2s)$  and  $j' = t - j + 1$ . Now  $p_{s,t}^* = (s - p_t, s - p_{t-1}, \dots, s - p_1)$  and

$$(4.49) \quad \lambda_{\nu, p_{s,t}^*} = 2^{st-|p|} \prod_{i=1}^t \frac{\Gamma[\frac{1}{2}(\nu + 1 - i) + s - p_{t-i+1}]}{\Gamma[\frac{1}{2}(\nu + 1 - i)]}.$$

Therefore combining all these things

$$(4.50) \quad \begin{aligned} & \lambda_{\nu, p_{s,t}^*} \sum_{q \in \mathcal{P}_{st-|p|}, \ell(q) \leq t} b_q y_q(A) = \lambda_{\nu, p_{s,t}^*} |A|^s y_p(A^{-1}) \\ &= \mathcal{E}_W\{|AW|^s y_p(A^{-1}W^{-1})\} \\ &= \sum_{q \in \mathcal{P}_{st-|p|}, \ell(q) \leq t} b_q \lambda_{\nu q} y_q(A). \end{aligned}$$

It follows that  $b_q=0$  if  $q \neq p_{s,t}^*$ . Therefore

$$(4.51) \quad |A|^s y_p(A^{-1}) = b_{p_{s,t}^*} y_{p_{s,t}^*}(A).$$

Putting  $A = I_t$  we obtain

$$(4.52) \quad b_{p_{s,t}^*} = y_p(I_t) / y_{p_{s,t}^*}(I_t).$$

■

The second distribution is a "multivariate F" distribution. There are many ways to generalize the univariate F distribution to the multivariate case. Here we work with the following version. For other generalizations see Johnson and Kotz(1972).

**Lemma 4.5.** *Let the columns of  $X_1 : k \times \nu_1$ ,  $X_2 : k \times \nu_2$  ( $\nu_2 > 2h(p) + k - 1$ ) be independently distributed according to  $N(O, \Sigma)$ . Let  $W = X_1'(X_2X_2')^{-1}X_1$ . Then*

$$(4.53) \quad \mathcal{E}_W\{y_p(AW)\} = \lambda_{kp} \prod_{i=1}^{\ell(p)} \frac{\Gamma[\frac{1}{2}(\nu_2 - k + i) - p_i]}{\Gamma[\frac{1}{2}(\nu_2 - k + i)]2^{p_i}} y_p(A).$$

*Proof:* Premultiplying  $X_1, X_2$  by  $\Sigma^{-\frac{1}{2}}$  we can take  $\Sigma = I_k$  without loss of generality. Then

$$(4.54) \quad \begin{aligned} \mathcal{E}_W y_p(AW) &= \mathcal{E}_{X_1} \mathcal{E}_{X_2} y_p(X_1 A X_1' (X_2 X_2')^{-1}) \\ &= \prod_{i=1}^{\ell(p)} \frac{\Gamma[\frac{1}{2}(\nu_2 - k + i) - p_i]}{\Gamma[\frac{1}{2}(\nu_2 - k + i)]2^{p_i}} \mathcal{E}_{X_1} y_p(A X_1' X_1) \\ &= \prod_{i=1}^{\ell(p)} \frac{\Gamma[\frac{1}{2}(\nu_2 - k + i) - p_i]}{\Gamma[\frac{1}{2}(\nu_2 - k + i)]2^{p_i}} \lambda_{kp} y_p(A). \end{aligned}$$

■

**Remark 4.3.** *It is more or less obvious to prove Lemma 4.5 for other definitions of multivariate F distribution.*

Our last distribution is multivariate beta distribution (Constantine(1963)). The following derivation is essentially the same as in Constantine(1963), but a little bit more

"probabilistic". Let  $W_1, W_2$  be independently distributed according to  $\mathcal{W}(\Sigma, \nu_1), \mathcal{W}(\Sigma, \nu_2)$  ( $\Sigma : k \times k$ ) respectively. Note that  $W = W_1 + W_2 \sim \mathcal{W}(\Sigma, \nu_1 + \nu_2)$ . Now the conditional density of  $W_1$  given  $W$  is

$$\begin{aligned}
 f(W_1 | W) &= c \frac{|W_1|^{\frac{\nu_1 - k - 1}{2}} \exp(-\frac{1}{2} \text{tr } \Sigma^{-1} W_1) |W - W_1|^{\frac{\nu_2 - k - 1}{2}} \exp(-\frac{1}{2} \text{tr } \Sigma^{-1} (W - W_1))}{|W|^{\frac{\nu_1 + \nu_2 - k - 1}{2}} \exp(-\frac{1}{2} \text{tr } \Sigma^{-1} W)} \\
 (4.55) \quad &= c \frac{|W_1|^{\frac{\nu_1 - k - 1}{2}} |W - W_1|^{\frac{\nu_2 - k - 1}{2}}}{|W|^{\frac{\nu_1 + \nu_2 - k - 1}{2}}},
 \end{aligned}$$

where

$$(4.56) \quad c = \frac{\prod_{i=1}^k \Gamma[\frac{1}{2}(\nu_1 + \nu_2 - i + 1)]}{\pi^{k(k-1)/4} \prod_{i=1}^k \Gamma[\frac{1}{2}(\nu_1 - i + 1)] \Gamma[\frac{1}{2}(\nu_2 - i + 1)]}.$$

Note that terms involving  $\Sigma$  cancel out in (4.55). Therefore the conditional distribution does not depend on  $\Sigma$ . When  $W = I$ ,  $f(W_1 | I)$  is called multivariate beta density:

$$(4.57) \quad f(W_1 | I) = c |W_1|^{\frac{\nu_1 - k - 1}{2}} |I - W_1|^{\frac{\nu_2 - k - 1}{2}}.$$

Since this density is orthogonally invariant the conditional distribution of  $W_1$  given  $W = I$  is orthogonally invariant. Now we want to evaluate  $c_p$  in

$$(4.58) \quad \mathcal{E}\{y_p(AW_1) | W = I\} = c_p y_p(A).$$

For a positive definite  $A$  let  $A^{\frac{1}{2}} = \Gamma D^{\frac{1}{2}} \Gamma'$  where  $\Gamma$  is orthogonal and  $D$  is diagonal in  $A = \Gamma D \Gamma'$ . Now the conditional distribution of  $A^{\frac{1}{2}} W_1 A^{\frac{1}{2}}$  given  $W = I$  is the same as the conditional distribution of  $W_1$  given  $W = A$ . This follows from the above mentioned fact that the conditional distribution does not depend on  $\Sigma$ . Therefore

$$(4.59) \quad \mathcal{E}\{y_p(AW_1) | W = I\} = \mathcal{E}\{y_p(W_1) | W = A\}.$$

Letting  $A = W_1 + W_2$  we obtain from (4.58) and (4.59)

$$(4.60) \quad \mathcal{E}\{y_p(W_1) | W_1 + W_2\} = c_p y_p(W_1 + W_2).$$

Now taking unconditional expectation we obtain

$$(4.61) \quad \lambda_{\nu_1 p} y_p(\Sigma) = c_p \lambda_{\nu_1 + \nu_2, p} y_p(\Sigma).$$

Hence  $c_p = \lambda_{\nu_1 p} / \lambda_{\nu_1 + \nu_2, p}$ . Now we have proved



**Lemma 4.6.** *Let  $W_1$  have the density (4.57). Then*

$$(4.62) \quad \mathcal{E}_{W_1} y_p(AW_1) = \frac{\lambda_{\nu_1 p}}{\lambda_{\nu_1 + \nu_2, p}} y_p(A).$$

Variations of the above three lemmas can be found in Khatri(1966), Subrahmaniam(1976).

#### §4.4 Coefficients of $u_q$ in $y_p$

In this section we study coefficients of  $u_q$ 's when zonal polynomials are expressed as linear combinations of  $u_q$ 's. For definiteness we work with  ${}_1a_{pq}$  in  ${}_1y_p = u_p + \sum {}_1a_{pq}u_q$ . If rank  $A=1,2$  all the relevant coefficients are known and we can compute  $y_p(A)$  explicitly. We review this first. After that we study several recurrence relations between the coefficients. When rank  $A > 2$  these recurrence relations are not enough to compute the values of zonal polynomials  $y_p(A)$  for all  $p$ . Nonetheless they seem to be very useful. Coefficients of  $M_q$ 's will be discussed in the next section and  $\tau_q$ 's in Section 4.6. We discuss relative advantages of various bases on the way.

##### 4.4.1 Rank 1 and rank 2 cases

If  $A$  is symmetric and rank  $A=1$  then  $A$  has only one nonzero root. Let  $A = \text{diag}(\alpha_1, 0, \dots, 0)$  without loss of generality. By Corollary 4.1  $y_p(A) = 0$  if  $\ell(p) \geq 2$ . Therefore only onepart partitions  $p = (p_1)$  count. Obviously

$$(4.63) \quad {}_1y_{(p_1)}(A) = u_{(p_1)}(A) = \alpha_1^{p_1}.$$

Therefore in this case zonal polynomials reduce to powers of  $\alpha_1$ .

Now suppose rank  $A=2$ . Let  $A = \text{diag}(\alpha_1, \alpha_2, 0, \dots, 0)$ . We have to consider only partitions with two parts  $p = (p_1, p_2)$ . Now we use Corollary 4.2:

$$(4.64) \quad {}_1y_{(p_1, p_2)}(A) = |A|^{p_2} {}_1y_{(p_1 - p_2)}(A),$$

where  $(p_1 - p_2)$  is a onepart partition. Therefore it suffices to know the value of a zonal polynomial of onepart partition evaluated at rank 2 matrix  $A$ . Actually zonal polynomials

of onepart partitions are known explicitly. If we let  $\beta_1 = 1, \beta_2 = \dots = \beta_k = 0$  in (3.82) we obtain

$$(4.65) \quad \prod_{i=1}^k (1 - 2\theta\alpha_i)^{-\frac{1}{2}} = \sum_{n=0}^{\infty} (\theta^n/n!) d_{(n)} Z_{(n)}(\mathbf{A}) {}_1b_{(n)}.$$

Now in (3.77)  $(\text{tr } C)^n = \mathcal{U}_{(n)}(C)$  and  $\mathcal{U}_{(n)}(C)$  term appears only in  $Z_{(n)}(C)$ . Therefore  $(\text{tr } C)^n = {}_1\mathcal{Y}_{(n)}(C) + \dots = (1/{}_1b_{(n)}) Z_{(n)}(C) + \dots$ . Hence  $d_{(n)} = {}_1b_{(n)}^{-1}$ . We have

$$(4.66) \quad \prod_{i=1}^k (1 - 2\theta\alpha_i)^{-\frac{1}{2}} = \sum_{n=0}^{\infty} (\theta^n/n!) Z_{(n)}(\mathbf{A}).$$

The left hand side can be expanded as follows.

$$(4.67) \quad \begin{aligned} \prod_{i=1}^k (1 - 2\theta\alpha_i)^{-\frac{1}{2}} &= \{1 - (2\theta u_1 - 4\theta^2 u_2 + \dots)\}^{-\frac{1}{2}} \\ &= \sum_{n=0}^{\infty} (2\theta)^n \sum_{p \in \mathcal{P}_n} \frac{1}{p_1!} \frac{1}{2} \frac{3}{2} \dots \frac{2p_1 - 1}{2} \binom{p_1}{p_1 - p_2, p_2 - p_3, \dots, p_{\ell(p)}} \\ &\quad \cdot (-1)^{(p_2 - p_3) + (p_4 - p_5) + \dots} \mathcal{U}_p. \end{aligned}$$

This follows from the fact that  $\mathcal{U}_p$  being a product of  $p_1$  terms comes only from  $p_1$ -th power term in the expansion of  $(1 - 2\theta u_1 + \dots)^{-1/2}$ . Comparing (4.66) and (4.67) we obtain

$$(4.68) \quad Z_{(n)} = 2^n n! \sum_{p \in \mathcal{P}_n} (-1)^{(p_2 - p_3) + (p_4 - p_5) + \dots} \frac{(\frac{1}{2})_{p_1}}{(p_1 - p_2)! \dots p_{\ell(p)}!} \mathcal{U}_p.$$

Note that  ${}_1b_{(n)} = 1 \cdot 3 \dots (2n - 1)$ ,  $2^n n! = 2 \cdot 4 \dots (2n)$ ,  $|A|^{p_2} \mathcal{U}_{(q_1, q_2)}(\mathbf{A}) = \mathcal{U}_{(q_1 + p_2, q_2 + p_2)}(\mathbf{A})$ .

Therefore if  $\text{rank } \mathbf{A} = 2$  we obtain from (4.64)

$$(4.69) \quad \begin{aligned} {}_1\mathcal{Y}_{(p_1, p_2)}(\mathbf{A}) &= \frac{2 \cdot 4 \dots (2p_1 - 2p_2)}{1 \cdot 3 \dots (2p_1 - 2p_2 - 1)} \\ &\quad \cdot \sum_{(q_1, q_2) \in \mathcal{P}_{p_1 - p_2}} (-1)^{q_2} \frac{(\frac{1}{2})_{q_1}}{(q_1 - q_2)! q_2!} \mathcal{U}_{(q_1 + p_2, q_2 + p_2)}(\mathbf{A}). \end{aligned}$$

See formula (130) in James(1964).

If  $\text{rank } \mathbf{A} = 3$  what we have to know is the values of zonal polynomials of twopart partitions evaluated at a rank 3 matrix  $\mathbf{A}$ . Obviously things become more and more complicated as  $\text{rank } \mathbf{A}$  increases. However several useful recurrence relations on the coefficients can be obtained.

## 4.4.2 Recurrence relations on the coefficients

We present here three recurrence relations. The first one is already used in deriving (4.69).

**Lemma 4.7.** *If  $k \geq \ell(p)$ ,  $k \geq \ell(q)$ , then*

$$(4.70) \quad {}_1a_{pq} = {}_1a_{p+(1^k), q+(1^k)}.$$

*Proof:* Let  $A$  be  $k \times k$ . Then

$$(4.71) \quad \begin{aligned} |A| {}_1y_p(A) &= |A| \{ \mathcal{U}_p(A) + \sum_{q < p, \ell(q) \leq k} {}_1a_{pq} \mathcal{U}_q(A) \} \\ &= \mathcal{U}_{p+(1^k)}(A) + \sum_{q < p, \ell(q) \leq k} {}_1a_{pq} \mathcal{U}_{q+(1^k)}(A). \end{aligned}$$

By Theorem 4.2

$$(4.72) \quad \begin{aligned} |A| {}_1y_p(A) &= {}_1y_{p+(1^k)}(A) \\ &= \mathcal{U}_{p+(1^k)}(A) + \sum_{q' < p+(1^k), \ell(q') \leq k} {}_1a_{p+(1^k), q'} \mathcal{U}_{q'}(A). \end{aligned}$$

Comparing (4.71) and (4.72) we obtain by Lemma 4.2  ${}_1a_{pq} = {}_1a_{p+(1^k), q+(1^k)}$ . ■

**Remark 4.4.** *Theorem 4.2 has been known and Lemma 4.7 is almost an immediate consequence. However it does not seem to have been explicitly stated.*

The next one is in a sense conjugate to Lemma 4.7. Let  $p = (p_1, \dots, p_\ell) \in \mathcal{P}_n$  and  $m \geq p_1 = h(p)$ . We denote by  $(m, p)$  the partition  $(m, p_1, p_2, \dots, p_\ell) \in \mathcal{P}_{n+m}$ .

**Theorem 4.4.** *Let  $m \geq h(p)$ . Then*

$$(4.73) \quad \frac{\partial^m}{\partial \alpha_{k+1}^m} {}_1y_{(m,p)}(\alpha_1, \dots, \alpha_{k+1}) = m! {}_1y_p(\alpha_1, \dots, \alpha_k).$$

*Proof:* Let

$$(4.74) \quad \begin{aligned} y_{(m,p)}(\alpha_1, \dots, \alpha_{k+1}) &= \sum_{q \leq (m,p), q \in \mathcal{P}_{n+m}} a_{(m,p), q} \mathcal{U}_q(\alpha_1, \dots, \alpha_{k+1}) \\ &= \sum_{(m, q') \leq (m,p), q' \in \mathcal{P}_n} a_{(m,p), (m, q')} \mathcal{U}_{(m, q')}(\alpha_1, \dots, \alpha_{k+1}) \\ &\quad + \sum_{q \in \mathcal{P}_{n+m}, h(q) < m} a_{(m,p), q} \mathcal{U}_q(\alpha_1, \dots, \alpha_{k+1}). \end{aligned}$$

We differentiate (4.74)  $m$  times with respect to  $\alpha_{k+1}$ . Now the degree of  $\alpha_{k+1}$  in

$$(4.75) \quad \mathcal{U}_q(\alpha_1, \dots, \alpha_{k+1}) = \left( \sum \alpha_i \right)^{q_1 - q_2} \left( \sum_{i < j} \alpha_i \alpha_j \right)^{q_2 - q_3} \dots$$

is  $q_1 = (q_1 - q_2) + (q_2 - q_3) + \dots + q_{\ell(q)}$ . Therefore the terms in the second summation on the right hand side of (4.74) drop out. Now  $\mathcal{U}_{(m, q')}(\alpha_1, \dots, \alpha_{k+1})$  is a product of  $m = (m - q'_1) + \dots + q'_{\ell(q')}$  elementary symmetric functions  $u_r(\alpha_1, \dots, \alpha_{k+1})$  which are linear in  $\alpha_{k+1}$ . Therefore differentiating  $\mathcal{U}_{(m, q')}$   $m$  times we are left with the term where each  $u_r$  is differentiated exactly once. Furthermore

$$(4.76) \quad \frac{\partial}{\partial \alpha_{k+1}} u_r(\alpha_1, \dots, \alpha_{k+1}) = u_{r-1}(\alpha_1, \dots, \alpha_k).$$

Therefore by the chain rule of differentiation

$$(4.77) \quad \begin{aligned} & \frac{\partial^m}{\partial \alpha_{k+1}^m} \mathcal{U}_{(m, q')}(\alpha_1, \dots, \alpha_{k+1}) \\ &= m! \left\{ \frac{\partial}{\partial \alpha_{k+1}} u_1(\alpha_1, \dots, \alpha_{k+1}) \right\}^{m - q'_1} \left\{ \frac{\partial}{\partial \alpha_{k+1}} u_2(\alpha_1, \dots, \alpha_{k+1}) \right\}^{q'_1 - q'_2} \dots \\ &= m! u_1(\alpha_1, \dots, \alpha_k)^{q'_1 - q'_2} u_2(\alpha_1, \dots, \alpha_k)^{q'_2 - q'_3} \dots \\ &= m! \mathcal{U}_{q'}(\alpha_1, \dots, \alpha_k). \end{aligned}$$

Let  $\tilde{\mathcal{Y}}(\alpha_1, \dots, \alpha_k) = (\partial^m / \partial \alpha_{k+1}^m) \mathcal{Y}_{(m, p)}(\alpha_1, \dots, \alpha_{k+1})$ . Then

$$(4.78) \quad \begin{aligned} \tilde{\mathcal{Y}}(\alpha_1, \dots, \alpha_k) &= \sum_{q \leq p} a_{(m, p), (m, q)} \frac{\partial^m}{\partial \alpha_{k+1}^m} \mathcal{U}_{(m, q)}(\alpha_1, \dots, \alpha_{k+1}) \\ &= m! \sum_{q \leq p} a_{(m, p), (m, q)} \mathcal{U}_q(\alpha_1, \dots, \alpha_k). \end{aligned}$$

We replaced  $q'$  by  $q$  and  $(m, q) \leq (m, p)$  by  $q \leq p$  since  $(m, q) \leq (m, p)$  if and only if  $q \leq p$ . We want to prove that  $\tilde{\mathcal{Y}}$  is a zonal polynomial. This can be done by showing that  $\tilde{\mathcal{Y}}$  satisfies the condition of Theorem 3.1.

Let  $A = \text{diag}(\alpha_1, \dots, \alpha_{k+1})$  and  $A_1 = \text{diag}(\alpha_1, \dots, \alpha_k)$ . Then exactly as above we obtain

$$(4.79) \quad \begin{aligned} \frac{\partial^m}{\partial \alpha_{k+1}^m} \mathcal{Y}_{(m, p)}(AW) &= m! \sum_{q \leq p} a_{(m, p), (m, q)} W(k+1)^{m - q_1} \left( \sum_{i_1}^k \alpha_{i_1} W(i_1, k+1) \right)^{q_1 - q_2} \dots \\ &\quad \cdot \left( \sum_{i_1 < \dots < i_{\ell(q)}}^k \alpha_{i_1} \dots \alpha_{i_{\ell(q)}} W(i_1, \dots, i_{\ell(q)}, k+1) \right)^{q_{\ell(q)}}. \end{aligned}$$

Let  $W$  be partitioned as

$$(4.80) \quad W = \begin{pmatrix} W_{11} & w_{k+1} \\ w'_{k+1} & w_{k+1,k+1} \end{pmatrix},$$

where  $w_{k+1,k+1} = W(k+1)$  is a scalar. Let

$$(4.81) \quad W_{11 \cdot k+1} = W_{11} - w_{k+1} w'_{k+1} / w_{k+1,k+1}.$$

Then by the well known identity on the determinant of partitioned matrices we have

$$(4.82) \quad W(i_1, \dots, i_r, k+1) = w_{k+1,k+1} W_{11 \cdot k+1}(i_1, \dots, i_r).$$

Therefore in (4.79)  $w_{k+1,k+1}^m = w_{k+1,k+1}^{(m-q_1)+(q_1-q_2)+\dots}$  comes out as a common factor and we obtain

$$(4.83) \quad \begin{aligned} \frac{\partial^m}{\partial \alpha_{k+1}^m} y_{(m,p)}(AW) &= m! \sum_{q \leq p} a_{(m,p),(m,q)} w_{k+1,k+1}^m \mathcal{U}_q(A_1 W_{11 \cdot k+1}) \\ &= w_{k+1,k+1}^m \tilde{y}(A_1 W_{11 \cdot k+1}) \end{aligned}$$

Now if  $W$  is distributed according to  $\mathcal{W}(I_{k+1}, \nu)$  then  $w_{k+1,k+1}$  and  $W_{11 \cdot k+1}$  are independently distributed according to  $\chi^2(\nu)$ ,  $\mathcal{W}(I_k, \nu - 1)$  respectively. (See Srivastava and Khatri(1979), Theorem 3.3.5 or Mardia, Kent, and Bibby(1979), Theorem 3.4.6.) Therefore taking expectation with respect to  $W$

$$(4.84) \quad \begin{aligned} \lambda_{\nu,(m,p)} \tilde{y}(A_1) &= \lambda_{\nu,(m,p)} \frac{\partial^m}{\partial \alpha_{k+1}^m} y_{(m,p)}(\alpha_1, \dots, \alpha_{k+1}) \\ &= \mathcal{E}_W \left\{ \frac{\partial^m}{\partial \alpha_{k+1}^m} y_{(m,p)}(AW) \right\} \\ &= \mathcal{E}_W \{ w_{k+1,k+1}^m \tilde{y}(A_1 W_{11 \cdot k+1}) \} \\ &= 2^m \frac{\Gamma(\frac{\nu}{2} + m)}{\Gamma(\frac{\nu}{2})} \mathcal{E}_W \{ \tilde{y}(A_1 W_{11 \cdot k+1}) \}. \end{aligned}$$

Hence

$$(4.85) \quad \mathcal{E}_W \tilde{y}(A_1 W_{11 \cdot k+1}) = \left\{ \lambda_{\nu,(m,p)} / 2^m \frac{\Gamma(\frac{\nu}{2} + m)}{\Gamma(\frac{\nu}{2})} \right\} \tilde{y}(A_1).$$

Now

$$(4.86) \quad \lambda_{\nu,(m,p)} = 2^{m+n} \frac{\Gamma(\frac{\nu}{2} + m)}{\Gamma(\frac{\nu}{2})} \prod_{i=1}^{\ell(p)} \frac{\Gamma(\frac{\nu-i}{2} + p_i)}{\Gamma(\frac{\nu-i}{2})}.$$

Therefore

$$(4.87) \quad \lambda_{\nu, (m, p)} / 2^m \frac{\Gamma(\frac{\nu}{2} + m)}{\Gamma(\frac{\nu}{2})} = \lambda_{\nu-1, p}.$$

By Theorem 3.1 we conclude

$$(4.88) \quad \tilde{y} = c y_p.$$

Comparing the leading coefficient we obtain  $c = m!$  in (4.73). ■

**Corollary 4.3.** *If  $m \geq h(p)$  then*

$$(4.89) \quad {}_1 a_{pq} = {}_1 a_{(m, p), (m, q)}.$$

*Proof:* From (4.73) and (4.77)

$$(4.90) \quad \begin{aligned} & m! (\mathcal{U}_p + \sum_{q < p} {}_1 a_{pq} \mathcal{U}_q) \\ &= m! {}_1 \mathcal{Y}_p = \frac{\partial^m}{\partial \alpha_{k+1}^m} {}_1 \mathcal{Y}_{(m, p)} \\ &= m! (\mathcal{U}_p + \sum_{q < p} {}_1 a_{(m, p), (m, q)} \mathcal{U}_q). \end{aligned}$$

Therefore (4.89) holds. ■

In terms of the diagram of  $p$  Lemma 4.7 corresponds to adding a column to the left of the diagram and Corollary 4.3 corresponds to adding a row to the top. In this sense they are “conjugate”. There might be a deeper reason for this conjugacy.

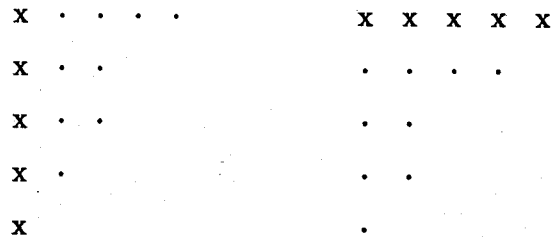


Figure 4.1.

Our third recurrence relation follows from Lemma 4.4.

**Lemma 4.8.** *If  $s \geq h(p), s \geq h(q), t \geq \ell(p), t \geq \ell(q)$ , then*

$$(4.91) \quad {}_1a_{pq} = {}_1a_{p^*,t,q^*,t}.$$

*Proof:* From Lemma 4.4

$$(4.92) \quad |A|^s {}_1y_p(A^{-1}) = c {}_1y_{p^*,t}(A),$$

or

$$(4.93) \quad \begin{aligned} |A|^s u_p(A^{-1}) + \sum_{q < p, \ell(q) \leq t} {}_1a_{pq} |A|^s u_q(A^{-1}) \\ = c(u_{p^*}(A) + \sum_{q' < p^*, \ell(q') \leq t} {}_1a_{p^*,q'} u_{q'}(A)), \end{aligned}$$

where  $p^* = p^*_{s,t}$  and  $c = {}_1y_p(I_t) / {}_1y_{p^*}(I_t)$ . Now let  $A = \text{diag}(\alpha_1, \dots, \alpha_t)$ . Then

$$(4.94) \quad \begin{aligned} |A| u_r(A^{-1}) &= (\alpha_1 \cdots \alpha_t) \sum_{i_1 < \dots < i_r}^t \frac{1}{\alpha_{i_1} \cdots \alpha_{i_r}} \\ &= \sum_{j_1 < \dots < j_{t-r}}^t \alpha_{j_1} \cdots \alpha_{j_{t-r}} \\ &= u_{t-r}(A). \end{aligned}$$

Note that (4.94) is true for  $r = 0, t$  if we define  $u_0 = 1$ . Therefore

$$(4.95) \quad \begin{aligned} |A|^s u_q(A^{-1}) &= |A|^{s-q_1} \{ |A|^{q_1} u_1(A^{-1})^{q_1-q_2} \cdots u_{\ell(q)}(A^{-1})^{q_{\ell(q)}} \} \\ &= u_t(A)^{s-q_1} u_{t-1}(A)^{q_1-q_2} \cdots u_{t-\ell(q)}(A)^{q_{\ell(q)}} \\ &= u_{q^*,t}(A), \end{aligned}$$

because  $q^*_{s,t} = (s, \dots, s, s - q_{\ell(q)}, \dots, s - q_2, s - q_1)$  and  $\ell(q^*_{s,t}) = t$ . Substituting (4.95) into (4.93) we obtain

$$(4.96) \quad \begin{aligned} u_{p^*}(A) + \sum_{q < p, \ell(q) \leq t} {}_1a_{pq} u_{q^*,t}(A) \\ = c \{ u_{p^*}(A) + \sum_{q' < p^*, \ell(q') \leq t} {}_1a_{p^*,q'} u_{q'}(A) \}. \end{aligned}$$

Therefore by Lemma 4.2  ${}_1a_{pq} = {}_1a_{p^*,t,q^*,t}$ ,  $c=1$ . ■

**Remark 4.5.** *Again this lemma is much easier to grasp in terms of the diagram. See Figure 2.2.*

Looking at Table 2 in Parkhurst and James(1974) we find that the above three recurrence relations give a lot of coefficients without any calculation (except that the table is for  $Z_p$  rather than for  ${}_1Y_p$ ). However it seems to be still a long way to obtain all coefficients along this line.

#### §4.5 Conversion to coefficients of $M_q$ and James' partial differential equation

So far we have been mainly working with  $\mathcal{U}_p$ 's. But in view of Lemma 2.5, Lemma 4.1 etc. we could have worked with  $M_p$ 's as well. We defined zonal polynomials in connection with the Wishart distribution and it was more straightforward to define zonal polynomials in terms of  $\mathcal{U}_p$ 's in that setting. But when it comes to obtaining coefficients it seems easier to work with  $M_p$ 's. In this section we translate every result in Section 4.4 into the coefficients of  $M_p$ 's. Another big advantage of working with monomial symmetric functions is a partial differential equation by James(1968), from which he derived a recurrence relation on the coefficients of monomial symmetric functions in a zonal polynomial. (Note that the recurrence relations of Section 4.4.2 were on the coefficients of  $\mathcal{U}_q$ 's in different zonal polynomials. Here the recurrence relation is on the coefficients in one zonal polynomial.) Actually it is possible to develop a whole theory of zonal polynomials from the partial differential equation. This is done in a recent book by Muirhead(1982) explicitly and illustratively. We discuss the partial differential equation and the recurrence relation in Section 4.5.4.

Furthermore Jacob Towber (personal communication) has recently obtained a combinatorial method for determining the coefficients. His method involves several steps of counting related to the diagram of a partition. At the moment the combinatorics involved seems to be too complicated to obtain an explicit formula for the coefficients, but it might be carried out.

From the above discussion we see that we have much more information on the coefficients of  $M_p$ 's than on the coefficients of  $\mathcal{U}_p$ 's. Therefore in a sense it is pointless



to work with  $\mathcal{U}_p$ 's any more. However from a computational point of view it is easier to compute  $\mathcal{U}_q$ 's once we obtain the characteristic roots and the characteristic equation of a matrix  $A$ . We simply multiply the elementary symmetric functions. In the case of  $\mathcal{M}_p$ 's we have to multiply the roots in all possible ways and sum them up. The relative advantages of  $\mathcal{M}_p$ 's and  $\mathcal{U}_p$ 's should be judged from this viewpoint too.

#### 4.5.1 Rank 1 and rank 2 cases

Let  $p = (n)$  be a onepart partition. To express  $Z_{(n)}$  in monomial symmetric functions we can use the integral representation by Kates. This was done by Kates(1980). Letting  $r = 1$  in (3.76) we obtain

$$(4.97) \quad UAU'(1) = \sum_{i=1}^k \alpha_i u_{1i}^2,$$

where  $A = \text{diag}(\alpha_1, \dots, \alpha_k)$  and  $u_{1i}, i = 1, \dots, k$ , are independent standard normal variables. Therefore by (3.70)

$$(4.98) \quad Z_{(n)} = \mathcal{E}\left(\sum_{i=1}^k \alpha_i u_{1i}^2\right)^n.$$

Now the coefficient of  $\alpha_1^{p_1} \dots \alpha_\ell^{p_\ell}$  on the right hand side is

$$(4.99) \quad \binom{n}{p_1, p_2, \dots, p_\ell} \mathcal{E}\{u_{11}^{2p_1} \dots u_{1\ell}^{2p_\ell}\} = \frac{n!}{p_1! \dots p_\ell!} \frac{(2p_1)!}{2^{p_1} p_1!} \dots \frac{(2p_\ell)!}{2^{p_\ell} p_\ell!} \\ = n! 2^{-n} \binom{2p_1}{p_1} \dots \binom{2p_\ell}{p_\ell}.$$

Therefore

$$(4.100) \quad Z_{(n)} = n! 2^{-n} \sum_{p \in \mathcal{P}_n} \mathcal{M}_p(A) \prod_{i=1}^{\ell(p)} \binom{2p_i}{p_i}.$$

This looks nicer than (4.68).  ${}_1\mathcal{Y}_p$  has the leading coefficient 1, so

$$(4.101) \quad {}_1\mathcal{Y}_{(n)} = \binom{2n}{n}^{-1} \sum_{p \in \mathcal{P}_n} \mathcal{M}_p \prod_{i=1}^{\ell(p)} \binom{2p_i}{p_i}.$$

Now let  $A = \text{diag}(\alpha_1, \alpha_2)$ . Then for  $q = (q_1, q_2)$  ( $q_1 \neq q_2$ )

$$\begin{aligned}
 |A|^k M_q(A) &= (\alpha_1 \alpha_2)^k (\alpha_1^{q_1} \alpha_2^{q_2} + \alpha_2^{q_1} \alpha_1^{q_2}) \\
 &= \alpha_1^{q_1+k} \alpha_2^{q_2+k} + \alpha_2^{q_1+k} \alpha_1^{q_2+k} \\
 &= M_{(q_1+k, q_2+k)}(A).
 \end{aligned}
 \tag{4.102}$$

(The equality of the extreme left and the extreme right hand sides holds for  $q_1 = q_2$  too).

Therefore from (4.64) we obtain

$${}_1\mathcal{Y}_{(p_1, p_2)}(\alpha_1, \alpha_2) = \binom{2p_1 - 2p_2}{p_1 - p_2}^{-1} \sum_{(q_1, q_2) \in \mathcal{P}_{p_1 - p_2}} \binom{2q_1}{q_1} \binom{2q_2}{q_2} M_{(q_1 + p_2, q_2 + p_2)}(\alpha_1, \alpha_2).
 \tag{4.103}$$

This takes care of rank 1 and rank 2 cases.

#### 4.5.2 Again on the generating function of zonal polynomials

To express  $\tau_p$  in terms of  $M_q$ 's we can simply expand it and count various monomial terms. Therefore it seems easier to express the right hand side of (3.85) in  $M_q$ 's than in  $\mathcal{U}_q$ 's. Then we decompose the resulting positive definite matrix as  $LL'$  where  $L$  is lower triangular with positive diagonal elements. The elements of  $L$  give the desired coefficients. The development on page 34 goes through in exactly the same way except that we order  $\{\tau_p, p \in \mathcal{P}_n\}$  according to the lexicographic ordering of the conjugate partition  $p'$  (see Remark 2.3). We do not repeat it here.

Rather we notice here the similarity between two generating functions (3.82) and (4.66). Let  $\gamma_1, \dots, \gamma_{k^2}$  denote the  $k^2$  numbers  $\alpha_i \beta_j$ ,  $i = 1, \dots, k$ ,  $j = 1, \dots, k$ . Let  $C = \text{diag}(\gamma_1, \dots, \gamma_{k^2})$ . Then from (3.82) and (4.66) we have

$$\sum_{n=0}^{\infty} (\theta^n / n!) \sum_{p \in \mathcal{P}_n} d_p Z_p(A) Z_p(B) = \sum_{n=0}^{\infty} (\theta^n / n!) Z_{(n)}(C).
 \tag{4.104}$$

Hence

$$\sum_{p \in \mathcal{P}_n} d_p Z_p(A) Z_p(B) = Z_{(n)}(C).
 \tag{4.105}$$

Now we can substitute (4.100) into the right hand side. Then it reduces to expressing  $M_p(C)$  as a sum of products  $M_q(A)M_{q'}(B)$ . This seems nicer than directly expanding the right hand side of (3.85).

Finally we prove that the coefficient of  $M_{(1^k)}$  in  $Z_p$ ,  $p \in \mathcal{P}_k$  is  $k!$ . This is stated in James(1968).

**Lemma 4.9.** *Let  $p \in \mathcal{P}_k$  and  $A = \text{diag}(\alpha_1, \dots, \alpha_k)$ . Then*

$$(4.106) \quad \frac{\partial^k}{\partial \alpha_1 \partial \alpha_2 \dots \partial \alpha_k} Z_p(A) = k!.$$

Hence the coefficient of  $M_{(1^k)}$  in  $Z_p$  is  $k!$ .

*Proof:*

$$(4.107) \quad \prod_{i,j}^k (1 - 2\theta \alpha_i \beta_j)^{-\frac{1}{2}} = \sum_{n=0}^{\infty} (\theta^n / n!) \sum_{p \in \mathcal{P}_n} d_p Z_p(A) Z_p(B).$$

Differentiating this by  $\alpha_1, \alpha_2, \dots, \alpha_k$  we obtain

$$(4.108) \quad \prod_{i,j}^k \left( \sum_{j=1}^k \frac{\theta \beta_j}{1 - 2\theta \alpha_i \beta_j} \right) \prod_{i,j}^k (1 - 2\theta \alpha_i \beta_j)^{-\frac{1}{2}} \\ = \sum_{n=0}^{\infty} (\theta^n / n!) \sum_{p \in \mathcal{P}_n} d_p Z_p(B) \frac{\partial^k}{\partial \alpha_1 \dots \partial \alpha_k} Z_p(A).$$

Now  $\theta \beta_j / (1 - 2\theta \alpha_i \beta_j) = \theta \beta_j + \text{higher order in } \theta$ . Hence

$$(4.109) \quad \prod_{i,j}^k \left( \sum_{j=1}^k \frac{\theta \beta_j}{1 - 2\theta \alpha_i \beta_j} \right) = \theta^k \left( \sum_{j=1}^k \beta_j \right)^k + \text{higher order in } \theta.$$

Comparing the coefficients of  $\theta^k$  we obtain

$$(4.110) \quad \left( \sum_{j=1}^k \beta_j \right)^k = \sum_{p \in \mathcal{P}_k} \frac{d_p}{k!} Z_p(B) \frac{\partial^k}{\partial \alpha_1 \dots \partial \alpha_k} Z_p(A).$$

But by (3.77)

$$(4.111) \quad \left( \sum_{j=1}^k \beta_j \right)^k = (\text{tr } B)^k = \sum_{p \in \mathcal{P}_k} d_p Z_p(B).$$

Comparing (4.110) and (4.111) we obtain

$$\frac{\partial^k}{\partial \alpha_1 \dots \partial \alpha_k} Z_p(A) = k!$$

## 4.5.3 Recurrence relations of Section 4.4.2

Here we work again with the normalization  ${}_1\mathcal{Y}_p$ . Let

$$(4.112) \quad {}_1\mathcal{Y}_p = M_p + \sum_{q < p} {}_1b_{pq} M_q.$$

**Lemma 4.10.** *If  $k \geq \ell(p)$ ,  $k \geq \ell(q)$ , then*

$$(4.113) \quad {}_1b_{pq} = {}_1b_{p+(1^k), q+(1^k)}.$$

*Proof:* Let  $A = \text{diag}(\alpha_1, \dots, \alpha_k)$  where  $k \geq \ell(p)$ . Then

$$(4.114) \quad \begin{aligned} |A| M_q(A) &= (\alpha_1 \cdots \alpha_k) \sum_{(i_1, \dots, i_\ell) \subset (1, \dots, k)} \alpha_{i_1}^{q_1} \alpha_{i_2}^{q_2} \cdots \alpha_{i_\ell}^{q_\ell} \\ &= \sum_{(i_1, \dots, i_\ell) \subset (1, \dots, k)} \alpha_{i_1}^{q_1+1} \alpha_{i_2}^{q_2+1} \cdots \\ &= M_{q+(1^k)}(A). \end{aligned}$$

Note that this equality does not hold for augmented monomial symmetric functions. The equality above holds because the summation is over distinguishable terms and  $\alpha_{i_1}^{q_1} \cdots \alpha_{i_\ell}^{q_\ell}$  is distinguishable from  $\alpha_{j_1}^{q_1} \cdots \alpha_{j_\ell}^{q_\ell}$  if and only if  $(\alpha_1 \cdots \alpha_k) \alpha_{i_1}^{q_1} \cdots \alpha_{i_\ell}^{q_\ell}$  is distinguishable from  $(\alpha_1 \cdots \alpha_k) \alpha_{j_1}^{q_1} \cdots \alpha_{j_\ell}^{q_\ell}$ . For augmented monomial functions refer to (2.15). Now the lemma can be proved just as Lemma 4.7 if we replace  $\mathcal{U}_p, \mathcal{U}_q, {}_1a_{pq}$  in (4.71) by  $M_p, M_q, {}_1b_{pq}$  respectively.

■

**Lemma 4.11.** *Let  $h(p) \leq m$ . Then*

$$(4.115) \quad {}_1b_{pq} = {}_1b_{(m,p), (m,q)}.$$

*Proof:* The degree of  $\alpha_{k+1}$  in  $M_q(\alpha_1, \dots, \alpha_{k+1})$  is  $h(q)$ . Hence if  $h(q) < m$ , then

$$(4.116) \quad \frac{\partial^m}{\partial \alpha_{k+1}^m} M_q(\alpha_1, \dots, \alpha_{k+1}) = 0.$$

If  $h(q) = m$  let  $q = (m, q')$ . Then clearly

$$(4.117) \quad \frac{\partial^m}{\partial \alpha_{k+1}^m} M_q(\alpha_1, \dots, \alpha_{k+1}) = m! M_{q'}(\alpha_1, \dots, \alpha_k).$$

(Again this equality does not hold for  $\mathcal{A}M_q$ .) Now (4.90) holds with  $M_p, M_q, {}_1b_{pq}$  replacing  $\mathcal{U}_p, \mathcal{U}_q, {}_1a_{pq}$  respectively. This proves the lemma. ■

**Lemma 4.12.** If  $h(p) \leq s, h(q) \leq s, \ell(p) \leq t, \ell(q) \leq t$ , then

$$(4.118) \quad {}_1b_{pq} = {}_1b_{p_{s,t}^*, q_{s,t}^*}.$$

*Proof:* Let  $A = \text{diag}(\alpha_1, \dots, \alpha_t)$ ,  $q = (q_1, \dots, q_\ell)$ ,  $q_1 \leq s, \ell \leq t$ . Then

$$(4.119) \quad \begin{aligned} |A|^s M_q(A^{-1}) &= (\alpha_1 \cdots \alpha_t)^s \sum_{(i_1, \dots, i_\ell) \subset (1, \dots, t)} \frac{1}{\alpha_{i_1}^{q_1} \cdots \alpha_{i_\ell}^{q_\ell}} \\ &= \sum_{(i_1, \dots, i_\ell) \subset (1, \dots, t)} \alpha_{i_1}^{s-q_1} \alpha_{i_2}^{s-q_2} \cdots \alpha_{i_\ell}^{s-q_\ell} \cdot \alpha_{j_1}^s \cdots \alpha_{j_{t-\ell}}^s \\ &= M_{q_{s,t}^*}(A), \end{aligned}$$

where  $1 \leq j_1, \dots, j_{t-\ell} \leq t$  are indices not included in  $(i_1, \dots, i_\ell)$  and  $q_{s,t}^* = (s, \dots, s, s - q_\ell, \dots, s - q_2, s - q_1)$ . (Again (4.119) does not hold for  $\mathcal{A}M_q$ ). Now (4.93), (4.95), (4.96) hold with  $M_p, M_q, {}_1b_{pq}$  replacing  $\mathcal{U}_p, \mathcal{U}_q, {}_1a_{pq}$  respectively. This proves the lemma. ■

We have shown that the recurrence relations of Section 4.4.2 hold in exactly the same way for the coefficients of  $\mathcal{U}_q$ 's as for the coefficients of  $M_q$ 's.

In the next section we discuss James' partial differential equation and a recurrence relation derived from it. The mathematical development will be somewhat sketchy.

#### 4.5.4 James' partial differential equation and recurrence relation

James(1968) derived a partial differential equation satisfied by a zonal polynomial from the fact that a zonal polynomial is an "eigenfunction of the Laplace-Beltrami operator." Let  $A = \text{diag}(\alpha_1, \dots, \alpha_k)$ ,  $p = (p_1, \dots, p_\ell) \in \mathcal{P}_n$ . Then his partial differential equation is

$$(4.120) \quad \sum_{i=1}^k \alpha_i^2 \frac{\partial^2}{\partial \alpha_i^2} y_p(A) + \sum_{i \neq j}^k \frac{\alpha_i^2}{\alpha_i - \alpha_j} \frac{\partial}{\partial \alpha_i} y_p(A) = \left( \sum_{i=1}^{\ell} p_i(p_i - i + k - 1) \right) y_p(A).$$

This might seem a little bit strange because it depends on the number of variables ( $k$  appears in the summation on the right hand side). Let

$$(4.121) \quad a_1(p) = \sum_{i=1}^{\ell} p_i(p_i - i).$$

Then the right hand side of (4.120) can be written as

$$(4.122) \quad a_1(p)y_p(A) + n(k-1)y_p(A), \quad n = |p|.$$

To get rid of  $n(k-1)y_p(A)$  we notice the fact that for any  $f \in V_n(A)$ ,  $\sum_{i=1}^k \alpha_i (\partial/\partial \alpha_i) f = nf$ . To prove this it suffices to show that  $\sum \alpha_i (\partial/\partial \alpha_i) M_p(A) = n M_p(A)$  for any  $p \in \mathcal{P}_n$ . Now this follows from

$$(4.123) \quad \sum \alpha_i \frac{\partial}{\partial \alpha_i} \alpha_1^{p_1} \alpha_2^{p_2} \cdots \alpha_\ell^{p_\ell} = (p_1 + \cdots + p_\ell) \alpha_1^{p_1} \cdots \alpha_\ell^{p_\ell} = n \alpha_1^{p_1} \cdots \alpha_\ell^{p_\ell}.$$

Therefore

$$(4.124) \quad (k-1)n y_p(A) = (k-1) \sum_{i=1}^k \alpha_i \frac{\partial}{\partial \alpha_i} y_p(A).$$

But we can write

$$(4.125) \quad (k-1) \sum_{i=1}^k \alpha_i \frac{\partial}{\partial \alpha_i} y_p(A) = \sum_{j=1}^k \sum_{i \neq j}^k \alpha_i \frac{\partial}{\partial \alpha_i} y_p(A).$$

Now subtracting (4.125) from both sides of (4.120) and using the relation  $\alpha_i^2/(\alpha_i - \alpha_j) - \alpha_i = \alpha_j/(\alpha_i - \alpha_j)$  we obtain

$$(4.126) \quad \sum_{i=1}^k \alpha_i^2 \frac{\partial^2}{\partial \alpha_i^2} y_p(A) + \sum_{i \neq j}^k \frac{\alpha_i \alpha_j}{\alpha_i - \alpha_j} \frac{\partial}{\partial \alpha_i} y_p(A) = a_1(p) y_p(A),$$

which does not involve  $k$  as a coefficient and is valid for any number of variables. (4.126) was derived by Sugiura(1973) in an elementary way. Because his exposition is clear and readable (except that there are complications like a higher order partial differential equation and differential equations for complex zonal polynomials) we do not derive it here. Let

$$(4.127) \quad y_p(A) = \sum_{q \leq p} b_{pq} M_q(A).$$

Substituting this into (4.126) we obtain

$$(4.128) \quad \sum_{q \leq p} b_{pq} \left\{ \sum_{i=1}^k \alpha_i^2 \frac{\partial^2}{\partial \alpha_i^2} + \sum_{i \neq j}^k \frac{\alpha_i \alpha_j}{\alpha_i - \alpha_j} \frac{\partial}{\partial \alpha_i} \right\} M_q(A) \\ = a_1(p) \sum_{q \leq p} b_{pq} M_q(A).$$

Now

$$(4.129) \quad \sum_{i=1}^k \alpha_i^2 \frac{\partial^2}{\partial \alpha_i^2} M_q(A), \quad \sum_{i \neq j}^k \frac{\alpha_i \alpha_j}{\alpha_i - \alpha_j} \frac{\partial}{\partial \alpha_i} M_q(A)$$

can be expressed as sums of monomial symmetric functions. Then comparing both sides of (4.128) we can determine the coefficients  $b_{pq}$ . It is hard to visualize what is going on here unless one works out some examples. Muirhead(1981) does that very carefully using (4.120) rather than (4.126). See James(1968) also. Therefore we only sketch the procedure here.

Let  $q = (q_1, \dots, q_\ell)$ . It is fairly straightforward to verify that

$$(4.130) \quad \sum_{i=1}^k \alpha_i^2 \frac{\partial^2}{\partial \alpha_i^2} M_q = \sum_{i=1}^{\ell} q_i(q_i - 1) M_q,$$

$$(4.131) \quad \sum_{i \neq j} \frac{\alpha_i \alpha_j}{\alpha_i - \alpha_j} \frac{\partial}{\partial \alpha_i} M_q = - \sum_{i=1}^{\ell} q_i(i - 1) M_q + \text{lower order terms}.$$

Adding (4.130) and (4.131) we obtain

$$(4.132) \quad \mathcal{D} M_q = a_1(q) M_q + \text{lower order terms},$$

where

$$\mathcal{D} = \sum_{i=1}^k \alpha_i^2 \frac{\partial^2}{\partial \alpha_i^2} + \sum_{i \neq j}^k \frac{\alpha_i \alpha_j}{\alpha_i - \alpha_j} \frac{\partial}{\partial \alpha_i}.$$

It is fortunate to get only lower order terms by the differential operation. It is this triangular nature of the differential operation that enables one to determine  $b_{pq}$  recursively starting from the (arbitrary) leading coefficient  $b_{pp}$ . If one works out "lower order terms" (which is not hard) one arrives at the following rule by James(1968):

$$(4.133) \quad b_{pq} = \sum_{q < q' \leq p} \frac{((q_i + r) - (q_j - r)) b_{pq'}}{a_1(p) - a_1(q)},$$

where  $q'$  is an unordered partition of the form  $q' = (q_1, \dots, q_i + r, \dots, q_j - r, \dots, q_{\ell(q)})$ , ( $1 \leq r \leq q_j$ ). The summation is over all  $(i, j, r)$  where  $i < j$ ,  $r \geq 1$  such that when the unordered partition  $q'$  is ordered we have  $q < q' \leq p$ .

Actually in view of Theorem 4.1 we have to consider only partitions  $q, q'$  which are majorized by  $p$ .

The advantage of this method is that it is self-contained. It gives all coefficients of a single zonal polynomial without computing others. Therefore it is by far the best method if one is interested in computing few zonal polynomials. On the other hand if one wants to compute many zonal polynomials then relying exclusively on this method seems to involve a great deal of redundant computations in view of recurrence relations of Section 4.5.3.

**Remark 4.6.** *Logically the recurrence relation (4.133) is not complete until one shows that the denominator  $a_1(p) - a_1(q)$  is never zero. James(1968) states that (4.133) gives rise to positive  $b_{pq}$ 's. Since the numerator  $(q_i + r) - (q_j - r) = (q_i - q_j) + 2r$  is positive he seems to claim that  $a_1(p) - a_1(q) > 0$  for all relevant pairs  $p, q$ . By Theorem 4.1 it is enough to prove*

$$(4.134) \quad a_1(p) - a_1(q) > 0 \quad \text{for } p \succ q.$$

*Then this ensures that (4.133) works for all cases and nonzero  $b_{pq}$ 's are positive. (4.134) can be easily proved using techniques from the theory of majorization. See Marshall and Olkin(1979). We do not go into this here.*

#### §4.6 Coefficients of $\tau_q$ in $Z_p$

In this section we study the coefficients of  $\tau_p$ . The normalization  $Z_p$  seems to be most advantageous. An important fact about the coefficients of  $\tau_p$  is their orthogonality. We derive this first. See formulas (117) and (118) in James(1964), and Problem 13.3.9 in Farrell(1976).

For  $p \in \mathcal{P}_n$  let

$$(4.135) \quad Z_p = \sum_{q \in \mathcal{P}_n} g_{pq} \tau_q.$$

Let  $Z = (Z_{(n)}, Z_{(n-1,1)}, \dots, Z_{(1^n)})'$ ,  $G = (g_{pq})$ . Then (4.135) can be expressed in a matrix form as

$$(4.136) \quad Z = G\tau.$$



Now we recall that the transition matrix  $F$  in  $\mathcal{T}=F\mathcal{U}$  is lower triangular ((2.37)). Substituting this into (4.136) we obtain

$$(4.137) \quad Z = GF\mathcal{U}.$$

But  $Z = E\mathcal{U}$ . Hence

$$(4.138) \quad G = EF^{-1}$$

Now (3.87) shows

$$(4.139) \quad E'DE = F'CF,$$

where  $C = \text{diag}(c_p, p \in \mathcal{P}_n)$  is obtained in (3.84) and  $D = \text{diag}(d_p, p \in \mathcal{P}_n)$  is known as (3.88). From (4.138) and (4.139) we obtain

$$(4.140) \quad G'DG = C.$$

Inverting this

$$(4.141) \quad GC^{-1}G' = D^{-1}.$$

Coordinatewise

$$(4.142) \quad \sum_p d_p g_{pq} g_{pq'} = \delta_{qq'} c_q, \quad (\text{column orthogonality}),$$

$$(4.143) \quad \sum_q g_{pq} g_{p'q} / c_q = \delta_{pp'} / d_p, \quad (\text{row orthogonality}),$$

where  $\delta_{pp'}$  is Kronecker's delta.

Actually  $c_q, q \in \mathcal{P}_n$  coincide with the elements of the first row of  $G$ . To see this let  $\beta_1 = 1, \beta_2 = \dots = \beta_n = 0$  in (3.85). Then clearly  $\tau_p(I_1) = 1$  for every  $p$  and we have  $Z_{(n)}(A) = \sum_p c_p \tau_p(A)$  (see (4.66)) and this proves the claim. Therefore (4.143) can be written alternatively as

$$(4.144) \quad \sum_q g_{pq} g_{p'q} / g_{(n),q} = \delta_{pp'} / d_p.$$

One obvious advantage of working with  $\tau_p$ 's is that the coefficient matrix is readily invertible. From (4.140)

$$(4.145) \quad G^{-1} = C^{-1} G' D.$$

Therefore once we express zonal polynomials in terms of  $\tau_p$ 's then it is easy to express  $\tau_p$ 's (and their linear combinations) in zonal polynomials.

(4.144) was used to compute zonal polynomials in Parkhurst and James(1974) as follows. (i)  $\mathcal{U}_p$ 's are expressed in  $\tau_p$ 's. (ii) They are Gram-Schmidt orthogonalized relative to the orthogonality relation (4.144) starting from the lowest partition  $(1^n)$  upwards. Because of the triangularity of  $\Xi$  this clearly results in zonal polynomials.

Some  $g_{pq}$ 's can be explicitly obtained using the fact  $Z_p(I_k) = \lambda_{kp}$ . We regard  $\lambda_{kp}$  as a function in  $k$ . Then by (3.21) it is a polynomial in  $k$  of degree  $|p| = n$ . Now since  $t_r(I_k) = k$  for any  $r$  we obtain  $\tau_p(I_k) = k^{p_1 - p_2} k^{p_2 - p_3} \dots = k^{p_1} = k^{h(p)}$ . Therefore putting  $I_k$  in (4.135) we obtain

$$(4.146) \quad \lambda_{kp} = \sum_{q \in \mathcal{P}_n} g_{pq} k^{h(q)}.$$

This uniquely determines  $g_{pq}$  for  $q = (n)$ ,  $q = (n-1, 1)$ ,  $q = (1^n)$  because these are the only partitions in  $\mathcal{P}_n$  with  $h(q) = n, n-1, 1$  respectively. Now the leading coefficient of  $\lambda_{kp}$  is 1, hence

$$(4.147) \quad g_{p, (n)} = 1.$$

(4.147) was originally used by James to determine the normalization  $Z_p$ . Now let us look at the coefficient of  $k^{n-1}$  in  $\lambda_{kp}$ . It is

$$(4.148) \quad \begin{aligned} & \sum_{i=1}^{\ell(p)} \{(-i+1) + (-i+3) + \dots + (-i+1+2p_i-2)\} \\ &= \sum_{i=1}^{\ell(p)} \frac{1}{2} p_i \{(-i+1) + (-i+1+2p_i-2)\} \\ &= \sum_{i=1}^{\ell(p)} p_i (p_i - i) \\ &= a_1(p). \end{aligned}$$

$a_1(p)$  already appeared in (4.121). Therefore

$$(4.149) \quad g_{p,(n-1,1)} = a_1(p).$$

This is mentioned in the introductory part of Parkhurst and James(1964) in a somewhat mysterious form

$$(4.150) \quad g_{p,(n-1,1)} = \sum_{i=1}^{\ell(p)} p_i(p_i - 1) - \frac{1}{2} \sum_{j=1}^{h(p)} p'_j(p'_j - 1),$$

where  $p' = (p'_1, \dots, p'_{h(p)})$  is the conjugate partition of  $p$ . (4.149) and (4.150) seem to agree.

Now the coefficient of  $k$  in  $\lambda_{kp}$  is

$$(4.151) \quad \{2 \cdot 4 \cdots (2p_1 - 2)\} \{(-1)(1) \cdots (2p_2 - 3)\} \cdots = 2^{n-1}(p_1 - 1)! \prod_{i=2}^{\ell(p)} \left(-\frac{i-1}{2}\right)_{p_i}.$$

Hence

$$(4.152) \quad g_{p,(1^n)} = 2^{n-1}(p_1 - 1)! \prod_{i=2}^{\ell(p)} \left(-\frac{i-1}{2}\right)_{p_i}.$$

This does not seem to have been noticed.

Now from (4.138)

$$(4.153) \quad \Xi = GF, \quad \Xi = (\xi_{pq}), \quad F = (f_{pq}).$$

$F$  is lower triangular. Therefore the last column of  $\Xi$  is  $f_{(1^n)(1^n)}$  times the last column of  $G$ . By (2.33)

$$(4.154) \quad \tau_{(1^n)} = t_n = (-1)^{n-1} n (\mathcal{U}_{(1^n)} + \cdots).$$

Hence  $f_{(1^n)(1^n)} = (-1)^{n-1} n$ . Therefore

$$(4.155) \quad \xi_{p,(1^n)} = f_{(1^n)(1^n)} g_{p,(1^n)} = (-2)^{n-1} n (p_1 - 1)! \prod_{i=2}^{\ell(p)} \left(-\frac{i-1}{2}\right)_{p_i}.$$

Making use of (4.145) gives another set of identities.

$$(4.156) \quad \mathcal{T} = G^{-1}Z = C^{-1}G'DZ.$$

Now  $DZ = (d_{(n)}Z_{(n)}, \dots, d_{(1^n)}Z_{(1^n)})'$  and  $d_p Z_p$  is denoted by  $C_p$  (3.89). Therefore

$$(4.157) \quad \tau = C^{-1}G'(C_{(n)}, \dots, C_{(1^n)})'.$$

Comparing the second element we obtain

$$(4.158) \quad \begin{aligned} \tau_{(n-1,1)} &= t_1^{n-2} t_2 \\ &= \frac{1}{c_{(n-1,1)}} \sum_{p \in \mathcal{P}_n} a_1(p) C_p \\ &= \frac{1}{n(n-1)} \sum_{p \in \mathcal{P}_n} a_1(p) C_p, \end{aligned}$$

where  $c_{(n-1,1)} = n(n-1)$  is given by (3.84). (4.158) was given by Sugiura and Fujikoshi(1959) by a different method. They derive more identities of this kind. Now looking at the last element we obtain

$$(4.159) \quad \begin{aligned} \tau_{(1^n)} &= t_n = M_{(n)} \\ &= \frac{1}{c_{(1^n)}} 2^{n-1} \sum_{p \in \mathcal{P}_n} C_p \cdot (p_1 - 1)! \prod_{i=2}^{\ell(p)} \left( -\frac{i-1}{2} \right)_{p_i} \\ &= \frac{1}{(n-1)!} \sum_{p \in \mathcal{P}_n} C_p \cdot (p_1 - 1)! \prod_{i=2}^{\ell(p)} \left( -\frac{i-1}{2} \right)_{p_i}. \end{aligned}$$

What are advantages and disadvantages of working with  $\tau_p$ 's? One advantage is that we do not have to compute characteristic roots of  $A$  to compute  $\tau_p(A)$ . (One only needs traces of powers of  $A$ .) Another advantage is the orthogonality discussed above. A serious drawback of  $\tau_p$  is that we have to compute  $\tau_p(A)$  for all  $p \in \mathcal{P}_n$  even if the rank of  $A$  is small. In usual statistical computations rank  $A$  is fixed and not too big. It is a covariance matrix for example. Since the number of partitions grows very fast as  $n$  increases if one wants to compute  $Z_p(A)$  for  $|p|$  big it seems better to use  $\mathcal{U}_q$ 's or  $M_q$ 's. The growth of the number of partitions  $p$  with  $\ell(p) \leq k$  ( $k$ : fixed) is much smaller than the growth of the number of all partitions. See Table 4.1 in David, Kendall, and Barton(1966).

#### §4.7 Variations of the integral representation of zonal polynomials

In this section we explore various variations of the integral representation (Theorem 3.7) discussed in Section 3.3. We first replace  $U$  by the  $k \times k$  uniform orthogonal matrix  $H$ .

**Theorem 4.5.** (James, 1979) For  $k \times k$  symmetric  $A$

$$(4.160) \quad \begin{aligned} \frac{y_p(A)}{y_p(I_k)} &= \mathcal{E}_H\{\Delta_1^{p_1-p_2} \dots \Delta_\ell^{p_\ell}\} \\ &= \mathcal{E}_H\left\{\prod_{i=1}^{\ell} [HAH'(1, \dots, i)]^{p_i-p_{i+1}}\right\}, \end{aligned}$$

where  $p = (p_1, \dots, p_\ell) \in \mathcal{P}_n$  and the  $k \times k$  orthogonal  $H$  is uniformly distributed.

*Proof:* As in Lemma 3.2 it is easy to check that  $\mathcal{E}_H\{\Delta_1^{p_1-p_2} \dots \Delta_\ell^{p_\ell}\} \in V_{n,k}(A)$ . Therefore we can write

$$(4.161) \quad \mathcal{E}_H\{\Delta_1^{p_1-p_2} \dots \Delta_\ell^{p_\ell}\} = \sum_{q \in \mathcal{P}_n, \ell(q) \leq k} a_q Z_q(A).$$

Replacing  $A$  by  $UAU'$  where  $k \times k$   $U$  is as in Theorem 3.7 and taking expectation with respect to  $U$  we obtain

$$(4.162) \quad Z_p(A) = \sum_{q \in \mathcal{P}_n, \ell(q) \leq k} a_q \lambda_{kp} Z_q(A).$$

This being true for any symmetric  $k \times k$   $A$  we conclude from Theorem 4.1

$$a_q = 0, q \neq p, a_p = \frac{1}{\lambda_{kp}} = \frac{1}{Z_p(I_k)}.$$

Since (4.160) is independent of normalization we can have  $y_p$  instead of  $Z_p$  in (4.160). ■

**Corollary 4.4.** (Kates) Let  $X: k \times k$  have an orthogonally biinvariant distribution then

$$(4.163) \quad \mathcal{E}_X\{\Delta_1^{p_1-p_2} \dots \Delta_\ell^{p_\ell}\} = \frac{y_p(A) \mathcal{E}_X\{y_p(X'X)\}}{\{y_p(I_k)\}^2}.$$

where  $\Delta_i = XAX'(1, \dots, i)$ .

*Proof:* We replace  $X$  by  $H_1XH_2$  where  $H_1$  and  $H_2$  are independently uniformly distributed. The distribution of  $X$  is unchanged. Now taking expectation with respect to  $H_1$  (Theorem 4.5) and  $H_2$  (Theorem 3.3) successively we obtain

$$(4.164) \quad \begin{aligned} \mathcal{E}_X\{\Delta_1^{p_1-p_2} \dots \Delta_\ell^{p_\ell}\} &= \mathcal{E}_{H_1, X, H_2}\left\{\prod_{i=1}^{\ell} [H_1XH_2AH_2'X'H_1'(1, \dots, i)]^{p_i-p_{i+1}}\right\} \\ &= \mathcal{E}_{X, H_2}\{y_p(XH_2AH_2'X')\}/y_p(I_k) \\ &= y_p(A) \mathcal{E}_X\{y_p(X'X)\}/\{y_p(I_k)\}^2. \end{aligned}$$

■

**Remark 4.7.** As in Remark 3.11  $X$  can be rectangular. If  $X$  is  $m \times k$ , then  $\{y_p(I_k)\}^2$  on the right hand side of (4.164) is replaced by  $y_p(I_k) y_p(I_m)$ .

An easy modification of the above formulas produces another set of identities.

**Theorem 4.6.** Let  $U_1, U_2$  be  $k \times k$  matrices whose entries are independent standard normal variables. Then for  $k \times k$   $A$

$$(4.165) \quad {}_1b_p Z_p(AA') = \varepsilon_{U_1, U_2} \left\{ \prod_{i=1}^{\ell} [U_1 A U_2(1, \dots, i)]^{2p_i - 2p_{i+1}} \right\},$$

where  ${}_1b_p$  is given by (4.39).

*Proof:* Let the singular value decomposition of  $A$  be  $A = \Gamma_1 D \Gamma_2$  where  $\Gamma_1, \Gamma_2$  are orthogonal,  $D = \text{diag}(\delta_1, \dots, \delta_k)$  and  $\delta_1^2, \dots, \delta_k^2$  are the characteristic roots of  $AA'$ . Since the order of  $\delta_1, \dots, \delta_k$  and the sign of each  $\delta_i$  can be arbitrary in the singular value decomposition we see that (4.165) is a homogeneous symmetric polynomial in  $\delta_1^2, \dots, \delta_k^2$ . Furthermore (4.165) does not change when  $A, U_1, U_2$  are augmented as in the proof of Theorem 3.7. Therefore to prove that (4.165) is a zonal polynomial it suffices to check whether it satisfies the condition of Theorem 3.1.

Now replace  $A$  by  $U_3 A$  where  $U_3$  is independent of  $U_1, U_2$ . Then

$$(4.166) \quad \begin{aligned} & \varepsilon_{U_1, U_2, U_3} \left\{ \prod_{i=1}^{\ell} [U_1 U_3 A U_2(1, \dots, i)]^{2p_i - 2p_{i+1}} \right\} \\ &= \varepsilon_{U_1, U_2, U_3} \left\{ \prod_{i=1}^{\ell} [U_3 U_1 A U_2(1, \dots, i)]^{2p_i - 2p_{i+1}} \right\} \\ &= \varepsilon_{U_1, U_2, T, H} \left\{ \prod_{i=1}^{\ell} [T H U_1 A U_2(1, \dots, i)]^{2p_i - 2p_{i+1}} \right\} \\ &= \lambda_{kp} \varepsilon_{U_1, U_2} \left\{ \prod_{i=1}^{\ell} [U_1 A U_2(1, \dots, i)]^{2p_i - 2p_{i+1}} \right\}, \end{aligned}$$

where  $T, H$  are as in Lemma 3.7. Now to obtain the normalizing constant we put  $A = I_k$ . Then

$$\begin{aligned}
{}_1b_p \lambda_{kp} &= \mathcal{E}_{U_1, U_2} \left\{ \prod_{i=1}^{\ell} [U_1 U_2(1, \dots, i)]^{2p_i - 2p_{i+1}} \right\} \\
(4.167) \quad &= \mathcal{E}_{T, U_2} \left\{ \prod_{i=1}^{\ell} [T U_2(1, \dots, i)]^{2p_i - 2p_{i+1}} \right\} \\
&= \lambda_{kp} \mathcal{E}_{U_2} \left\{ \prod_{i=1}^{\ell} U_2(1, \dots, i)^{2p_i - 2p_{i+1}} \right\}.
\end{aligned}$$

Hence  ${}_1b_p = \mathcal{E}_U \{ \prod_{i=1}^{\ell} U(1, \dots, i)^{2p_i - 2p_{i+1}} \}$ . Now consider (3.70) and (3.76). Then we see that the coefficient of  $\alpha_1^{p_1} \dots \alpha_{\ell}^{p_{\ell}}$  is given just by  $\mathcal{E}_U \{ \prod_{i=1}^{\ell} U(1, \dots, i)^{2p_i - 2p_{i+1}} \}$ . Therefore it is the leading coefficient of  $Z_p$  and is given by (4.33).

**Corollary 4.5.**

$$(4.168) \quad \frac{{}_1b_p Z_p(AA')}{Z_p(I_k)} = \mathcal{E}_{H_1, U_2} \left\{ \prod_{i=1}^{\ell} [H_1 A U_2(1, \dots, i)]^{2p_i - 2p_{i+1}} \right\}.$$

$$(4.169) \quad \frac{{}_1b_p Z_p(AA')}{\{Z_p(I_k)\}^2} = \mathcal{E}_{H_1, H_2} \left\{ \prod_{i=1}^{\ell} [H_1 A H_2(1, \dots, i)]^{2p_i - 2p_{i+1}} \right\}.$$

*Proof:* (4.168) and (4.169) can be proved successively as in the proof of Theorem 4.5. ■

**Corollary 4.6.** *Let  $X_1, X_2$  be independent and have orthogonally biinvariant distributions. Then*

$$\begin{aligned}
(4.170) \quad &Z_p(AA') \frac{{}_1b_p \mathcal{E}_{X_1} \{ Z_p(X_1 X_1') \} \mathcal{E}_{X_2} \{ Z_p(X_2 X_2') \}}{\{Z_p(I_k)\}^4} \\
&= \mathcal{E}_{X_1, X_2} \left\{ \prod_{i=1}^{\ell} [X_1 A X_2(1, \dots, i)]^{2p_i - 2p_{i+1}} \right\}.
\end{aligned}$$

*Proof:* Replace  $X_1$  by  $H_1 X_1 H_3$  and  $X_2$  by  $H_4 X_2 H_2$ . Then taking expectation with respect to  $H_1, H_2, H_3, H_4$  successively we obtain (4.170). ■

**Remark 4.8.** *Generalization to rectangular matrices is straightforward.*

# Complex zonal polynomials

In this chapter we study complex zonal polynomials, i.e. zonal polynomials associated with the complex normal and the complex Wishart distributions. The complex multivariate normal distribution is used in the frequency analysis of multiple time series and complex zonal polynomials are useful for noncentral distributions arising in this setting. Other than that the practical applicability of complex zonal polynomials seems rather limited. Actually our main reason of studying them is that they are *simpler* than real zonal polynomials. If one compares Farrell(1980) and Chapter 1 of Macdonald(1979) it becomes apparent that complex zonal polynomials are the same as homogeneous symmetric polynomials called the Schur functions and the latter have been extensively studied. We will make this connection clear. Hopefully developing complex zonal polynomials gives further insights into the real case.

The theory of the complex normal and the Wishart distributions very closely parallels that of the real case (see Goodman(1963) or Brillinger(1975)) and it turns out that our development of Chapter 3 and Chapter 4 can be directly translated into the complex case. It seems customary to put a  $\sim$  to denote corresponding objects in the complex case. For example we use  $\tilde{Z}_p, \tilde{C}_p, \tilde{Y}_p$ , etc. With this convention the translation of the results in Chapter 3 and 4 are almost immediate.

## §5.1 The complex normal and the complex Wishart distributions.

We give a brief summary of the complex normal and the complex Wishart dis-



tributions. Let  $x, y$  be independently distributed according to  $\mathcal{N}(0, 1/2)$  and let  $z = x + iy$ .  $z$  is said to have the *standard complex normal distribution*. Or we say that  $z$  is a *standard complex normal (random) variable*. Now let  $A$  be an  $n \times n$  matrix with complex elements and let

$$(5.1) \quad \mathbf{u} = (u_1, \dots, u_k)' = A(z_1, \dots, z_k)',$$

where  $z_1, \dots, z_k$  are independent standard complex normal variables. This scheme generates a family of distributions called the multivariate complex normal distribution. Its density (with respect to  $\prod_1^k d(\Re u_i) \prod_1^k d(\Im u_i)$ ) is given by

$$(5.2) \quad f(\mathbf{u}) = \frac{1}{\pi^k |\Sigma|} \exp(-\mathbf{u}^* \Sigma^{-1} \mathbf{u})$$

where  $*$  means conjugate transpose and  $\Sigma = \mathcal{E} \mathbf{u} \mathbf{u}^* = A A^*$ . If  $\mathbf{u}$  has the density (5.2) we denote this by  $\mathbf{u} \sim \mathcal{CN}(\mathbf{0}, \Sigma)$ . Now suppose that  $\mathbf{u}_1, \dots, \mathbf{u}_n$  are independently distributed according to  $\mathcal{CN}(\mathbf{0}, \Sigma)$ . Let  $\tilde{W} = \sum_{i=1}^n \mathbf{u}_i \mathbf{u}_i^*$ . The distribution of  $\tilde{W}$  is called the complex Wishart distribution and its density (with respect to  $\prod_{i=1}^k d\tilde{w}_{ii} \prod_{i < j} d(\Re \tilde{w}_{ij}) d(\Im \tilde{w}_{ij})$ ) is given by

$$(5.3) \quad f(\tilde{w}) = \frac{|\tilde{w}|^{n-k} \exp(-\text{tr } \Sigma^{-1} \tilde{w})}{\pi^{p(p-1)/2} \prod_{i=1}^k \Gamma(n-i+1) |\Sigma|^n}$$

This distribution is denoted by  $\mathcal{CW}(\Sigma, n)$ .

Let  $\tilde{W} = \tilde{T} \tilde{T}^*$  be the (unique) triangular decomposition of a positive definite Hermitian matrix where  $\tilde{T} = (\tilde{t}_{ij})$  is a lower triangular matrix with positive diagonal elements. Analogous to Lemma 3.3 we have the following lemma.

**Lemma 5.1.** *Let  $\tilde{W}$  be distributed according to  $\mathcal{CW}(I_k, \nu)$ . Let  $\tilde{W} = \tilde{T} \tilde{T}^*$ . Then  $\tilde{t}_{ij}$ ,  $i \geq j$ , are independently distributed.  $\sqrt{2} \tilde{t}_{ii} \sim \chi(2(\nu - i + 1))$  and  $\tilde{t}_{ij}$ ,  $i > j$ , are standard complex normal variables.*

*Proof:* See Goodman(1963), formula(1.8) ■

**Remark 5.1.**  $\tilde{t}_{ii}^2$  has the gamma density  $f(x) = (1/\Gamma(\nu - i + 1))x^{\nu-i}e^{-x}$ .

With this lemma we are ready to translate the results of Chapter 3 and 4.

## §5.2 Derivation and properties of complex zonal polynomials

For ease of comparison of the results here and the results of Chapter 3 and 4 we will consistently put  $\sim$  on corresponding objects of the complex case. This sometimes results in somewhat unnatural notation, for example if  $H$  is orthogonal then  $\tilde{H}$  is unitary etc. So much for the notation and now let us follow the development of real zonal polynomials step by step for a while. All proofs will be omitted since they are the same for the real and the complex cases.

We consider the following transformation.

$$(5.4) \quad (\tilde{\tau}_\nu \mathcal{U}_p)(\tilde{A}) = \mathcal{E}_{\tilde{W}}\{\mathcal{U}_p(\tilde{A}\tilde{W})\},$$

where  $\tilde{A}$  is Hermitian and  $\tilde{W} \sim \mathcal{CW}(I_k, \nu)$ .

**Lemma 5.2.** (corresponding to Lemma 3.2) If  $\tilde{A}$  is Hermitian, then

$$(5.5) \quad (\tilde{\tau}_\nu \mathcal{U}_p)(\tilde{A}) \in V_n(\tilde{A}).$$

**Corollary 5.1.** (Corollary 3.1)

$$(5.6) \quad (\tilde{\tau}_\nu \mathcal{U}_p)(\tilde{A}) = \tilde{\lambda}_{\nu p} \mathcal{U}_p(\tilde{A}) + \sum_{q < p} \tilde{a}_{pq} \mathcal{U}_q(\tilde{A}),$$

where  $\tilde{A}$  is Hermitian.

**Corollary 5.2.** (Corollary 3.2)

$$(5.7) \quad \begin{aligned} \tilde{\lambda}_{\nu p} &= \prod_{i=1}^{\ell(p)} \frac{\Gamma(p_i + \nu + 1 - i)}{\Gamma(\nu + 1 - i)} \\ &= \prod_{i=1}^{\ell(p)} (\nu + 1 - i)_{p_i} \\ &= \nu(\nu + 1) \cdots (\nu + p_1 - 1) \\ &\quad \cdot (\nu - 1) \nu \cdots (\nu - 1 + p_2 - 1) \\ &\quad \cdots \\ &\quad \cdot (\nu - \ell + 1) \cdots (\nu - \ell + p_\ell) \end{aligned}$$

where  $\ell = \ell(p)$  and  $(a)_k = a(a+1)\cdots(a+k-1)$ .

Corollary 5.1 shows that

$$(5.8) \quad \tilde{\tau}_\nu(\mathcal{U}) = \tilde{T}_\nu \mathcal{U},$$

where  $\tilde{T}_\nu$  is an upper triangular matrix with diagonal elements  $\tilde{t}_{pp} = \tilde{\lambda}_{\nu p}$ .

**Lemma 5.3.** *(Lemma 3.4) There exists an upper triangular matrix  $\tilde{\mathbf{E}}$  such that*

$$(5.9) \quad \tilde{\mathbf{E}}\tilde{T}_\nu = \tilde{\Lambda}_\nu\tilde{\mathbf{E}} \quad \text{for all } \nu.$$

where  $\tilde{\Lambda}_\nu = \text{diag}(\tilde{\lambda}_{\nu p}, p \in \mathcal{P}_n)$ .  $\tilde{\mathbf{E}}$  is uniquely determined up to a (possibly different) multiplicative constant for each row.

Using this  $\tilde{\mathbf{E}}$  we define complex zonal polynomials.

**Definition 5.1** *Complex zonal polynomials*

Let  $\tilde{\mathbf{E}}$  be as in Lemma 5.3. Complex zonal polynomials  $\tilde{y}_p$ ,  $p \in \mathcal{P}_n$  are defined by

$$(5.10) \quad \tilde{\mathbf{y}} = \begin{pmatrix} \tilde{y}_{(n)} \\ \tilde{y}_{(n-1,1)} \\ \cdot \\ \cdot \\ \tilde{y}_{(1^n)} \end{pmatrix} = \tilde{\mathbf{E}}\mathcal{U}.$$

Lemma 5.3 is a consequence of the fact that there exists  $\nu_0$  for which  $\tilde{\lambda}_{\nu_0 p}$ ,  $p \in \mathcal{P}_n$  are all different and the following lemma.

**Lemma 5.4.** *(Lemma 3.5)*

$$(5.11) \quad \tilde{T}_\nu \tilde{T}_\mu = \tilde{T}_\mu \tilde{T}_\nu.$$

We summarize these results in the following theorem.

**Theorem 5.1.** (Theorem 3.1) Let  $\tilde{Y}_p$  be a complex zonal polynomial then

$$(5.12) \quad \varepsilon_{\tilde{W}} \tilde{Y}_p(\tilde{A}\tilde{W}) = \tilde{\lambda}_{\nu p} \tilde{Y}_p(\tilde{A}),$$

where  $\tilde{W} \sim \mathcal{CW}(I_k, \nu)$ ,  $\tilde{A}$  is Hermitian, and  $\tilde{\lambda}_{\nu p}$  is given by (5.7). Conversely (5.12) (for all sufficiently large  $\nu$ ) implies that  $\tilde{Y}_p$  is a complex zonal polynomial.

Now we explore various integral identities satisfied by complex zonal polynomials.

**Definition 5.2** A random unitary matrix  $\tilde{H}$  is said to have the *Haar invariant distribution* or the *uniform distribution* if the distribution of  $\tilde{H}\tilde{\Gamma}$  is the same for every unitary  $\tilde{\Gamma}$ .

More formally

**Definition 5.2'** A probability measure  $P$  on the Borel field of unitary matrices is *Haar invariant* if

$$(5.13) \quad P(A) = P(A\tilde{\Gamma}),$$

for every unitary  $\tilde{\Gamma}$  and every Borel set  $A$ .

Existence and uniqueness of the Haar invariant distribution are established by the following two lemmas.

**Lemma 5.5.** (Lemma 3.6) Let two probability measures  $P_1, P_2$  satisfy (5.13). Then  $P_1(A) = P_2(A)$  for every Borel set  $A$ . Furthermore  $P_1(A) = P_1(A^*)$  where  $A^* = \{\tilde{H}^* \mid \tilde{H} \in A\}$ .

**Lemma 5.6.** (Lemma 3.7) Let  $\tilde{U} = (\tilde{u}_{ij})$  be a  $k \times k$  matrix such that  $\tilde{u}_{ij}$  are independent standard complex normal variables. Then with probability 1  $\tilde{U}$  can be uniquely expressed as

$$(5.14) \quad \tilde{U} = \tilde{T}\tilde{H},$$

where  $\tilde{T} = (\tilde{t}_{ij})$  is lower triangular with positive diagonal elements and  $\tilde{H}$  is unitary. Furthermore (i)  $\tilde{T}, \tilde{H}$  are independent, (ii)  $\tilde{H}$  is uniform, (iii)  $\tilde{t}_{ij}$  are all independent and  $\sqrt{2}\tilde{t}_{ii} \sim \chi(2(k-i+1))$   $\tilde{t}_{ij}, i > j, \sim \mathcal{CN}(0, 1)$ .

Now we obtain the "splitting property" of complex zonal polynomials.

**Theorem 5.2.** (Theorem 3.2) Let  $\tilde{A}, \tilde{B}$  be  $k \times k$  Hermitian matrices. Then

$$(5.15) \quad \mathcal{E}_{\tilde{H}} \tilde{y}_p(\tilde{A} \tilde{H} \tilde{B} \tilde{H}^*) = \tilde{y}_p(\tilde{A}) \tilde{y}_p(\tilde{B}) / \tilde{y}_p(I_k),$$

where  $k \times k$  unitary  $\tilde{H}$  has the uniform distribution.

**Definition 5.3** A random Hermitian matrix  $\tilde{V}$  is said to have a *unitarily invariant distribution* if for every unitary  $\tilde{I}$ ,  $\tilde{I} \tilde{V} \tilde{I}^*$  has the same distribution as  $\tilde{V}$ .

As in the real case Theorem 5.1 generalizes to unitarily invariant distributions.

**Theorem 5.3.** (Theorem 3.3) Suppose that  $\tilde{V}$  has a unitarily invariant distribution, then for Hermitian  $\tilde{A}$

$$(5.16) \quad \mathcal{E}_{\tilde{V}} \tilde{y}_p(\tilde{A} \tilde{V}) = c_p \tilde{y}_p(\tilde{A}),$$

where

$$(5.17) \quad c_p = \mathcal{E}_{\tilde{V}} \{ \tilde{y}_p(\tilde{V}) \} / \tilde{y}_p(I_k).$$

Unitarily invariant distributions are characterized as follows.

**Lemma 5.7.** (Lemma 3.8) Let  $\tilde{V} = \tilde{H} \tilde{D} \tilde{H}^*$  where  $\tilde{H}$  is unitary and  $\tilde{D}$  is diagonal. Let  $\tilde{H}$  and  $\tilde{D}$  be independently distributed such that  $\tilde{H}$  has the uniform distribution. (Diagonal elements of  $\tilde{D}$  can have any distribution.) Then  $\tilde{V}$  has a unitarily invariant distribution. Conversely all unitarily invariant distributions can be obtained in this way.

We can replace  $\tilde{H}$  in Theorem 5.2 by  $\tilde{U}$  whose elements are independent standard complex normal variables.

**Theorem 5.4.** (Theorem 3.4) Let  $\tilde{U} = (\tilde{u}_{ij})$  be a  $k \times k$  matrix such that  $\tilde{u}_{ij}$  are independent standard complex normal variables. Then for Hermitian  $\tilde{A}, \tilde{B}$

$$(5.18) \quad \mathcal{E}_{\tilde{U}} \tilde{y}_p(\tilde{A} \tilde{U} \tilde{B} \tilde{U}^*) = \frac{\tilde{\lambda}_{kp}}{\tilde{y}_p(I_k)} \tilde{y}_p(\tilde{A}) \tilde{y}_p(\tilde{B}).$$

As in the real case this leads to the following observation.

**Theorem 5.5.** (Theorem 3.5)  $\tilde{b}_p = \tilde{\lambda}_{kp}/\tilde{y}_p(I_k)$  is a constant independent of  $k$ .

Unitarily biinvariant distributions are defined in an obvious way.

**Definition 5.4** A random matrix  $\tilde{X}$  has a *unitarily biinvariant distribution* if for every unitary  $\tilde{I}_1, \tilde{I}_2$ , the distribution of  $\tilde{I}_1 \tilde{X} \tilde{I}_2$  is the same as the distribution of  $\tilde{X}$ .

Now Theorem 5.2 and Theorem 5.4 generalize as follows.

**Theorem 5.6.** (Theorem 3.6) If  $\tilde{X}$  has a unitarily biinvariant distribution then for Hermitian  $\tilde{A}, \tilde{B}$

$$(5.19) \quad \mathcal{E}_{\tilde{X}} \tilde{y}_p(\tilde{A} \tilde{X} \tilde{B} \tilde{X}^*) = \gamma_p \tilde{y}_p(\tilde{A}) \tilde{y}_p(\tilde{B}),$$

where

$$(5.20) \quad \gamma_p = \mathcal{E}_{\tilde{X}} \{\tilde{y}_p(\tilde{X} \tilde{X}^*)\} / \{\tilde{y}_p(I_k)\}^2.$$

Characterization of unitarily biinvariant distributions can be given in an obvious way.

**Lemma 5.8.** (Lemma 3.9) Let  $\tilde{X} = \tilde{H}_1 \tilde{D} \tilde{H}_2$  where  $\tilde{H}_1, \tilde{H}_2$  are unitary and  $\tilde{D}$  is diagonal. Let  $\tilde{H}_1, \tilde{H}_2, \tilde{D}$  be independently distributed such that  $\tilde{H}_1, \tilde{H}_2$  have the uniform distribution. ( $\tilde{D}$  can have any distribution.) Then  $\tilde{X}$  has a unitarily biinvariant distribution. Conversely all unitarily biinvariant distributions can be obtained in this way.

**Remark 5.2.** The notion of unitarily biinvariant distributions applies to rectangular matrices as well.

Now we take a look at the integral representation of zonal polynomials in the complex case.

**Definition 5.5** A particular normalization of a zonal polynomial denoted by  $\tilde{Z}_p$  is defined by

$$(5.21) \quad \tilde{Z}_p(I_k) = \tilde{\lambda}_{kp},$$

or  $\tilde{b}_p = 1$  in Theorem 5.5.

**Theorem 5.7.** (Theorem 3.7) Let  $p = (p_1, \dots, p_\ell)$ . For  $k \times k$  Hermitian  $\tilde{A}$

$$(5.22) \quad \tilde{Z}_p(\tilde{A}) = \mathcal{E}_{\tilde{U}}\{\tilde{\Delta}_1^{p_1-p_2} \tilde{\Delta}_2^{p_2-p_3} \dots \tilde{\Delta}_\ell^{p_\ell}\},$$

where  $\tilde{\Delta}_i = \tilde{U}\tilde{A}\tilde{U}^*(1, \dots, i)$  is the  $i \times i$  upper left minor of  $\tilde{U}\tilde{A}\tilde{U}^*$  and  $\tilde{U}$  is a  $k \times k$  random matrix whose entries are independent standard complex normal variables.

(5.22) implies that  $\tilde{Z}_p(\tilde{A})$  is positive for positive definite  $\tilde{A}$  and increasing in each root. Furthermore using the Gale-Ryser theorem (see Remark 4.1 and Remark 4.2) the coefficients of  $M_q$  in  $\tilde{Z}_p$  are nonnegative and they are positive iff  $p \succ q$ .

As in the real case  ${}_1\tilde{b}_p$  denotes the leading coefficient of  $\tilde{Z}_p$ , namely

$$(5.23) \quad \tilde{Z}_p = {}_1\tilde{b}_p {}_1\tilde{y}_p.$$

**Theorem 5.8.** (Theorem 4.3)

$$(5.24) \quad \begin{aligned} {}_1\tilde{b}_p &= \prod_{i=1}^{\ell(p)} \prod_{j=1}^i (i-j+1+p_j-p_i)_{p_i-p_{i+1}} \\ &= \frac{\prod_{i=1}^{\ell(p)} (p_i - i + \ell(p))!}{\prod_{i < j} (p_i - p_j - i + j)}. \end{aligned}$$

Other than mentioning Theorem 4.3 we will not follow the development of Chapter 4. Of course all the results of Chapter 4 can be translated into the complex case as has been done so far. However, it is pointless to go into numerical aspects of complex zonal polynomials because, as mentioned above, complex zonal polynomials are the Schur functions and the Schur functions are already well studied. Although the translation of the results in Chapter 4 presents an alternative “probabilistic” derivation of properties of Schur functions, it is hardly more advantageous than a well developed standard approach to the subject. We will present the approach in Macdonald(1979) in the next section. The link between our approach so far and the one in Macdonald(1979) is given by Saw’s generating function.

Saw’s generating function in the complex case was introduced by Farrell(1980). Let  $\tilde{u}_{ij}$  be a standard complex normal variable. Then  $2|\tilde{u}_{ij}|^2 = 2\tilde{u}_{ij}\tilde{u}_{ij}^* \sim \chi^2(2)$  (i.e.  $|\tilde{u}_{ij}|^2$  has the standard exponential distribution). Therefore by considering  $\mathcal{E}_{\tilde{U}}\{\exp(\theta \operatorname{tr} \tilde{A}\tilde{U}\tilde{B}\tilde{U}^*)\}$  where  $\tilde{A} = \operatorname{diag}(\alpha_1, \dots, \alpha_k)$ ,  $\tilde{B} = \operatorname{diag}(\beta_1, \dots, \beta_k)$ , and  $\tilde{U}$  is composed of independent standard complex normal variables, we obtain

**Theorem 5.9.** (Theorem 3.8)

$$(5.25) \quad \prod_{i,j}^k (1 - \theta \alpha_i \beta_j)^{-1} = \sum_{n=0}^{\infty} (\theta^n / n!) \sum_{p \in \mathcal{P}_n} \tilde{d}_p \tilde{Z}_p(\tilde{A}) \tilde{Z}_p(\tilde{B}),$$

where  $\tilde{d}_p$  is determined by

$$(5.26) \quad (\text{tr } \tilde{A})^n = \sum_{p \in \mathcal{P}_n} \tilde{d}_p \tilde{Z}_p(\tilde{A}).$$

Coefficients of  $\tilde{Z}_p$  can be obtained as in the real case, namely (i) compare the coefficients of  $\theta^n$  in both sides of (5.25), (ii) express the left hand side as a quadratic form in  $M_p$  or  $\mathcal{U}_p$ , (iii) do the triangular decomposition to the resulting positive definite symmetric matrix of coefficients. Now it will be shown in the next section that Schur functions  $S_p$  satisfy the same generating function (5.25) and  $S_p$  is a linear combination of lower order  $M_q$ 's ( $S_p = \sum_{q \leq p} a_{pq} M_q$ ). Therefore Schur functions agree with the complex zonal polynomials by the uniqueness of the triangular decomposition of a positive definite symmetric matrix.

### §5.3 Schur functions and their determinantal expressions.

In this section we present the development in Chapter 1 of Macdonald(1979) leading to Saw's generating function. (It is in Section 1.4, page 33 in Macdonald's book.) This part of the book is entirely elementary and forms an excellent introduction into the theory of symmetric functions. In addition to several definitions we need only two lemmas to derive Saw's generating function. In order to be self-contained we give (detailed) proofs of them following Macdonald(1979).

Let  $p = (p_1, \dots, p_\ell) \in \mathcal{P}_n$ . The Schur function  $S_p(x_1, \dots, x_k)$  ( $k \geq \ell$ ) is defined by

$$(5.27) \quad S_p(x_1, \dots, x_k) = \det(x_j^{p_i+k-i})_{1 \leq i,j \leq k} / \det(x_j^{k-i})_{1 \leq i,j \leq k}$$

$$= \begin{vmatrix} x_1^{p_1+k-1} & \dots & x_k^{p_1+k-1} \\ x_1^{p_2+k-2} & \dots & x_k^{p_2+k-2} \\ \vdots & & \vdots \\ x_1^{p_k} & \dots & x_k^{p_k} \end{vmatrix} \div \begin{vmatrix} x_1^{k-1} & \dots & x_k^{k-1} \\ \vdots & & \vdots \\ x_1 & \dots & x_k \\ 1 & \dots & 1 \end{vmatrix}.$$



This is formula (35) in James(1964). If  $k < \ell$  we define  $S_p(x_1, \dots, x_k) = 0$ . Note that the denominator is the *Vandermonde determinant*

$$(5.28) \quad \det(x_j^{k-i}) = \prod_{i < j} (x_i - x_j).$$

Clearly the numerator has  $(x_i - x_j)$  as a factor because if  $x_i = x_j$  then  $\det(x_j^{p_i+k-i}) = 0$ . Running  $(i, j)$  over all pairs we see that the numerator has the Vandermonde determinant as a factor. Furthermore if  $x_i$  and  $x_j$  are interchanged then both the numerator and the denominator change the sign and the ratio remains the same. Therefore  $S_p(x_1, \dots, x_k)$  is a symmetric polynomial in  $x_i$ 's. It is easy to see that it is homogeneous of degree  $|p|$ . Now we want to show that

$$(5.29) \quad S_p(x_1, \dots, x_k, 0) = S_p(x_1, \dots, x_k).$$

The last column of  $(x_j^{p_i+k+1-i})_{1 \leq i, j \leq k+1}$  is

$$(x_{k+1}^{p_1+k}, \dots, x_{k+1}^{p_k+k}, x_{k+1}^{p_{k+1}})'$$

If  $x_{k+1} = 0$  it reduces to  $(0, \dots, 0, 1)'$ . (Note that  $p_{k+1} = 0$  by definition.) Hence if  $x_{k+1} = 0$  then  $\det(x_j^{p_i+k+1-i}) = \det(x_j^{p_i+k-i})$ , the right hand side being the  $k \times k$  principal minor of the matrix on the left hand side. Similarly  $\det(x_j^{k+1-i}) = \det(x_j^{k-i})$ . Therefore we have (5.29) and in general by induction

$$(5.30) \quad S_p(x_1, \dots, x_k, 0, \dots, 0) = S_p(x_1, \dots, x_k).$$

This shows that  $S_p \in V_n$ . Now let us look at the highest monomial in  $S_p$  of the form  $ax_1^{q_1} \dots x_k^{q_k}$   $((q_1, \dots, q_k) \in \mathcal{P}_n)$ . In  $\det(x_j^{p_i+k-i})$  and  $\det(x_j^{k-i})$  the similar terms are obtained by the products of the diagonal elements. They are

$$x_1^{p_1+k-1} x_2^{p_2+k-2} \dots x_k^{p_k}, \quad x_1^{k-1} \dots x_{k-1}$$

respectively. From  $S_p(x_1, \dots, x_k) \det(x_j^{k-i}) = \det(x_j^{p_i+k-i})$  we obtain

$$(ax_1^{q_1} \dots x_k^{q_k})(x_1^{k-1} \dots x_{k-1}) = x_1^{p_1+k-1} \dots x_k^{p_k}$$

Therefore  $a = 1$  and  $q = (q_1, \dots, q_k) = (p_1, \dots, p_k) = p$ . We summarize these results in a lemma.

**Lemma 5.9.**

$$(5.31) \quad S_p = M_p + \sum_{q < p} a_{pq} M_q,$$

and  $\{S_p, p \in \mathcal{P}_n\}$  forms a basis of  $V_n$ .

Now we introduce the  $r$ -th complete symmetric function  $h_r$  which is the coefficient of  $s^r$  in

$$(5.32) \quad H(s) = \prod_{i=1}^k (1 - sx_i)^{-1}.$$

For convenience we let  $h_0 = 1, h_r = 0$  for  $r < 0$ . Similarly  $u_0 = 1, u_r = 0$  for  $r < 0$ . From (2.28) and (5.32) we have

$$(5.33) \quad H(s)U(-s) = 1,$$

or equivalently

$$(5.34) \quad \sum_{r=0}^n (-1)^r u_r h_{n-r} = 0, \quad n = 1, 2, \dots$$

Recursively solving (5.34) we can express  $u_r$ 's (and their products) in terms of  $h_r$ 's and vice versa. Defining

$$(5.35) \quad \mathcal{H}_p = h_{p_1} h_{p_2} \cdots h_{p_\ell}, \quad p \in \mathcal{P}_n,$$

we see that  $\{\mathcal{H}_p, p \in \mathcal{P}_n\}$  forms still another basis of  $V_n$ . We also note that the relation between  $u_r$ 's and  $h_r$ 's corresponds to the relation between coefficients of autoregressive and moving average representations in time series.

**Lemma 5.10.** ((3.4) in Macdonald, (36) in James(1964))

$$(5.36) \quad \begin{aligned} S_p &= \det(h_{p_i - i + j})_{1 \leq i, j \leq k} \\ &= \begin{vmatrix} h_{p_1} & h_{p_1+1} & \cdots & h_{p_1+k-1} \\ h_{p_2-1} & h_{p_2} & \cdots & h_{p_2+k-2} \\ \vdots & \vdots & & \vdots \\ h_{p_k-k+1} & h_{p_k-k+2} & \cdots & h_{p_k} \end{vmatrix}. \end{aligned}$$

where  $k \geq \ell(p)$ .

*Proof:* We work with  $k$  variables  $x_1, \dots, x_k$ . Let  $u_r^{(j)}$  denote the elementary symmetric function of  $x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_k$  (omitting  $x_j$ ) and let  $M$  be the  $k \times k$  matrix

$$(5.37) \quad M = ((-1)^{k-i} u_{k-i}^{(j)}).$$

For  $p \in \mathcal{P}_n$ ,  $\ell(p) \leq k$  let  $k \times k$  matrices  $A_p, H_p$  be defined by

$$(5.38) \quad A_p = (x_j^{p_i}), \quad H_p = (h_{p_i-k+j}).$$

Now

$$U^{(j)}(s) = \sum_{r=1}^{k-1} u_r^{(j)} s^r = \prod_{i \neq j} (1 + s x_i).$$

Therefore

$$\begin{aligned} H(s)U^{(j)}(s) &= (1 - s x_j)^{-1} \\ &= \sum_{r=0}^{\infty} s^r x_j^r. \end{aligned}$$

Equating the coefficient of  $s^{p_i}$  we obtain

$$\sum_{r=1}^k h_{p_i-k+r} (-1)^{k-r} u_{k-r}^{(j)} = x_j^{p_i}.$$

Therefore

$$(5.39) \quad H_p M = A_p.$$

Taking the determinants we have

$$(5.40) \quad |H_p| |M| = |A_p| \quad \text{for any } p \in \mathcal{P}_n.$$

For  $p = \delta = (k-1, k-2, \dots, 1, 0)$ ,  $H_p$  is an upper triangular matrix with diagonal elements  $1 = h_0$ . Therefore  $|H_\delta| = 1$  and

$$|M| = |A_\delta| = \det(x_j^{k-i}),$$

which is the Vandermonde determinant. Therefore

$$(5.41) \quad |H_p| |A_\delta| = |A_p|.$$

Now replacing  $p$  by  $p + \delta$  we obtain

$$|A_{p+\delta}| = \det(x_j^{p_i+k-i}),$$

which is the numerator of (5.27). Furthermore

$$\begin{aligned} |H_{p+\delta}| &= \det(h_{p_i+k-i-k+j}) \\ &= \det(h_{p_i-i+j}). \end{aligned}$$

Hence

$$\begin{aligned} |H_{p+\delta}| &= |A_{p+\delta}|/|A_\delta| \\ &= \det(x_j^{p_i+k-i})/\det(x_j^{k-i}) \\ &= S_p. \end{aligned}$$

■

**Lemma 5.11.** ((4.3) in Macdonald)

$$(5.42) \quad \prod_{i,j}^k (1 - \theta x_i y_j)^{-1} = \sum_{n=0}^{\infty} \theta^n \sum_{p \in \mathcal{P}_n} S_p(x_1, \dots, x_k) S_p(y_1, \dots, y_k).$$

*Proof:* Replacing  $x_i$  by  $\theta x_i$  we can assume  $\theta = 1$  without loss of generality. Now from (5.32)

$$\begin{aligned} (5.43) \quad \prod_{i,j}^k (1 - x_i y_j)^{-1} &= \prod_{j=1}^k H(y_j) \\ &= \prod_j \left( \sum_{r=0}^{\infty} h_r(x) y_j^r \right) \\ &= \sum_{(a_1, \dots, a_k) \in \mathbb{N}^k} h_{a_1}(x) \cdots h_{a_k}(x) y_1^{a_1} \cdots y_k^{a_k} \\ &= \sum_p \mathcal{H}_p(x) \mathcal{M}_p(y), \end{aligned}$$

where  $(a_1, \dots, a_k)$  run over all  $k$ -tuples of nonnegative integers and  $p$  runs over all partitions (of length  $\leq k$ ). Denoting permutations by  $\pi$

$$\begin{aligned}
 & |A_\delta(x)||A_\delta(y)| \prod_{i,j}^k (1 - x_i y_j)^{-1} \\
 &= |A_\delta(x)| \left( \sum_{\pi} \text{sign}(\pi) \prod_{i=1}^k y_i^{k-\pi_i} \right) \sum_{(a_1, \dots, a_k) \in \mathbb{N}^k} h_{a_1} \dots h_{a_k} y_1^{a_1} \dots y_k^{a_k} \\
 (5.44) \quad &= |A_\delta(x)| \sum_{(a_1, \dots, a_k) \in \mathbb{N}^k} \sum_{\pi} \text{sign}(\pi) h_{a_1} \dots h_{a_k} y_1^{a_1+k-\pi_1} \dots y_k^{a_k+k-\pi_k} \\
 &= |A_\delta(x)| \sum_{(b_1, \dots, b_k) \in \mathbb{N}^k} \left( \sum_{\pi} \text{sign}(\pi) \prod_{i=1}^k h_{b_i-k+\pi_i} \right) \prod_{i=1}^k y_i^{b_i} \\
 &= |A_\delta(x)| \sum_{(b_1, \dots, b_k) \in \mathbb{N}^k} \det(h_{b_i-k+j}) \prod_{i=1}^k y_i^{b_i}.
 \end{aligned}$$

Now note that if  $b_i = b_j$ , then  $\det(h_{b_i-k+j}) = 0$ . Hence we can assume that  $b_i$ 's are all distinct in the summation. Suppose  $b_{\pi_1} > b_{\pi_2} > \dots$ , then

$$\det(h_{b_i-k+j}) = \det(h_{b_{\pi_i}-k+j}) \text{sign}(\pi).$$

Hence using (5.41) we obtain

$$\begin{aligned}
 & |A_\delta(x)| \sum_{(b_1, \dots, b_k) \in \mathbb{N}^k} \det(h_{b_i-k+j}) \prod_{i=1}^k y_i^{b_i} \\
 (5.45) \quad &= |A_\delta(x)| \sum_{b_1 > \dots > b_k \geq 0} \det(h_{b_i-k+j}) \sum_{\pi} \text{sign}(\pi) \prod_{i=1}^k y_i^{b_{\pi_i}} \\
 &= \sum_{b_1 > \dots > b_k} |A_{b_1, \dots, b_k}(x)| |A_{b_1, \dots, b_k}(y)|
 \end{aligned}$$

Now the set of all  $k$ -tuples  $(b_1, \dots, b_k)$  such that  $b_1 > \dots > b_k \geq 0$  agrees with the set  $\{p + \delta \mid p : \text{partition}, \ell(p) \leq k\}$ . Therefore

$$\begin{aligned}
 (5.46) \quad & |A_\delta(x)||A_\delta(y)| \prod_{i,j}^k (1 - x_i y_j)^{-1} = \sum_{b_1 > \dots > b_k} |A_{b_1, \dots, b_k}(x)| |A_{b_1, \dots, b_k}(y)| \\
 &= \sum_{p: \ell(p) \leq k} |A_{p+\delta}(x)| |A_{p+\delta}(y)|.
 \end{aligned}$$

This proves the lemma.  $\blacksquare$

Comparing (5.25) and (5.42) we have

$$(5.47) \quad \sum_{p \in \mathcal{P}_n} (\tilde{d}_p/n!) \tilde{Z}_p(A) \tilde{Z}_p(B) = \sum_{z \in \mathcal{P}_n} S_p(A) S_p(B) \\ = \sum_{p \in \mathcal{P}_n} {}_1\tilde{y}_p(A) {}_1\tilde{y}_p(B),$$

Hence  $S_p = {}_1\tilde{y}_p$ . Furthermore

$$(5.48) \quad \tilde{d}_p {}_1\tilde{b}_p^2 = n!, \quad n = |p|.$$

This was mentioned at the end of Section 4.3.

**Remark 5.3.** *There are two more determinantal expressions stated in James(1964). One which involves elementary symmetric functions (formula(37) in James(1964)) is found in (2.9') of Macdonald. Formula (38) in James(1964) is not given in Macdonald.*

From the viewpoint of numerical computation these determinantal expressions seem to be all we have to know. For a given matrix we calculate the roots or complete symmetric functions and evaluate the determinant, which can be easily done by computer. We do not have to know the coefficients of  $M_p$  or  $\mathcal{U}_p$  etc. to evaluate the Schur function. It might be worthwhile to look for an analogue of this for the real zonal polynomial. Another possibility is to express the real zonal polynomials in terms of  $\mathcal{X}_p$ 's. Still another possibility is to express the real zonal polynomials in terms of the Schur functions.

#### §5.4 Relation between the real and the complex zonal polynomials.

We finish this chapter by discussing some results which we were unable to derive by our elementary approach. James(1964) gives the following formula relating the complex and the real zonal polynomials:

$$(5.49) \quad \frac{Z_p(\mathbf{X}\mathbf{X}')}{Z_p(I_k)} = \varepsilon_H S_{2p}(\mathbf{X}\mathbf{H}),$$

where the  $k \times k$   $\mathbf{H}$  has the uniform distribution of orthogonal matrices,  $p = (p_1, \dots, p_l) \in \mathcal{P}_n$ , and  $2p = (2p_1, \dots, 2p_l) \in \mathcal{P}_{2n}$ . (Formula (34) in James(1964).) Furthermore he states

$$(5.50) \quad \varepsilon_H S_p(\mathbf{X}\mathbf{H}) = 0,$$

if one or more parts of  $p$  is odd. (Formula (40)). See also Theorem 12.11.6 and Remark 12.11.11 in Farrell (1976).

Given these results we can evaluate  $d_p$  in (3.77) as follows. First note that by replacing  $H$  by  $U$  where  $U$  is composed of independent standard (real) normal variables we obtain

$$\begin{aligned} \mathcal{E}_U S_{2p}(XU) &= \mathcal{E}_{T,H} S_{2p}(XTH) \\ &= \mathcal{E}_T Z_p(XTT'X')/Z_p(I_k) \\ &= Z_p(XX'). \end{aligned} \quad (5.51)$$

By (5.26) and (5.48)

$$\begin{aligned} (tr A)^{2n} &= \sum_{p \in \mathcal{P}_{2n}} \tilde{d}_p \tilde{Z}_p(A) \\ &= \sum_{p \in \mathcal{P}_{2n}} \tilde{d}_p {}_1\tilde{b}_p {}_1\tilde{y}_p(A) \\ &= \sum_{p \in \mathcal{P}_{2n}} (2n)! {}_1\tilde{b}_p^{-1} {}_1\tilde{y}_p(A). \end{aligned} \quad (5.52)$$

Now let  $A = \text{diag}(\alpha_1, \dots, \alpha_k)$  and replace  $A$  by  $AU$ . In this case

$$tr AU = \sum_{i=1}^k \alpha_i u_{ii} \sim \mathcal{N}(0, \sum \alpha_i^2).$$

Hence

$$\begin{aligned} \mathcal{E}_U (tr AU)^{2n} &= 1 \cdot 3 \cdots (2n-1) (\sum \alpha_i^2)^n \\ &= \frac{(2n)!}{2^n n!} (tr AA')^n. \end{aligned} \quad (5.53)$$

On the other hand by (5.52) and (5.51)

$$\mathcal{E}_U \sum_{p \in \mathcal{P}_{2n}} (2n)! {}_1\tilde{b}_p^{-1} {}_1\tilde{y}_p(AU) = \sum_{p \in \mathcal{P}_n} (2n)! {}_1\tilde{b}_{2p}^{-1} Z_p(AA'). \quad (5.54)$$

Hence comparing (5.53) and (5.54) we obtain

$$(tr AA')^n = \sum_{p \in \mathcal{P}_n} {}_1\tilde{b}_{2p}^{-1} 2^n n! Z_p(AA')$$

or

$$d_p = 2^n n! {}_1\tilde{b}_{2p}^{-1}. \quad (5.55)$$

Now (5.24) gives (3.88).

## Appendix: List of new results

### *Chapter 3:*

Proof of Lemma 3.2.

Definition 3.1.

Lemma 3.5.

Theorem 3.3.

Lemma 3.8.

Theorem 3.5.

Theorem 3.6.

Remark 3.11.

Proof of Theorem 3.7.

### *Chapter 4:*

Theorem 4.1.

Proof of Theorem 4.2.

Proof of Theorem 4.3.

Proof of Lemma 4.4.

Proof of Lemma 4.6.

Lemma 4.7.

Theorem 4.4.

Corollary 4.3.

Lemma 4.8.

Formula (4.105).

Proof of Lemma 4.9.

Lemma 4.10.

Lemma 4.11.

Lemma 4.12.

Formula (4.149).

Formula (4.152).

Formula (4.155).

Formula (4.159).

Theorem 4.6.

Corollary 4.5.

Corollary 4.6.

### *Chapter 5:*

All results in Section 5.2 which correspond to the new results in Chapter 3 and 4.

Derivation of (5.55) given (5.49) and (5.50).



## References

- Anderson, T.W. (1946). The noncentral Wishart distribution and certain problems of multivariate statistics. *Ann.Math.Statist.*, **17**, 409-431.
- Anderson, T.W. (1958). *An introduction to multivariate statistical analysis*. Wiley, New York.
- Anderson, T.W. and Girshick, M.A. (1944). Some extensions of the Wishart distribution. *Ann.Math.Statist.*, **15**, 345-357.
- Brillinger, D.R. (1975). *Time series: data analysis and theory*. Holt, Rinehart and Winston, New York.
- Constantine, A.G. (1963). Some noncentral distribution problems in multivariate analysis. *Ann. Math. Statist.*, **34**, 1270-1285.
- Constantine, A.G. (1966). The distribution of Hotelling's generalized  $T_0^2$ . *Ann. Math. Statist.*, **37**, 215-225.
- David, F.N., Kendall, M.G., and Barton, D.E. (1966). *Symmetric function and allied tables*. Cambridge University Press.
- Farrell, R.H. (1976) *Techniques of multivariate calculation*. Vol.520, Lecture Notes in Mathematics, Springer-Verlag, Berlin.
- Farrell, R.H. (1980). Calculation of complex zonal polynomials. In *Multivariate Analysis-V* (P.R.Krixhnaiah, ed.), North-Holland, pp. 301-320.

- Gantmacher, F.R. (1959). *The theory of matrices, Vol. I*. Chelsea, New York.
- Goodman, N.R. (1963). Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction). *Ann. Math. Statist.*, **34**, 152-177.
- Halmos, P.R. (1974). *Measure Theory*. Springer, New York.
- Herz, C.S. (1955). Bessel functions of matrix argument. *Ann. Math.*, **61**, 474-523.
- James, A.T. (1954). Normal multivariate analysis and the orthogonal group. *Ann. Math. Statist.*, **25**, 40-75.
- James, A.T. (1955a). The noncentral Wishart distribution. *Proc. Roy. Soc. (London), A*, **229**, 364-366.
- James, A.T. (1955b). A generating function for averages over the orthogonal group. *Proc. Roy. Soc. (London), A*, **229**, 367-375.
- James, A.T. (1960). The distribution of the latent roots of the covariance matrix. *Ann. Math. Statist.*, **31**, 151-158.
- James, A.T. (1961a). Zonal polynomials of the real positive definite symmetric matrices. *Ann. Math.*, **74**, 456-469.
- James, A.T. (1961b). The distribution of noncentral means with known covariances. *Ann. Math. Statist.*, **32**, 874-882.
- James, A.T. (1964). Distribution of matrix variates and latent roots derived from normal samples. *Ann. Math. Statist.*, **35**, 475-501.
- James, A.T. (1968). Calculation of zonal polynomial coefficients by use of the Laplace-Beltrami operator. *Ann. Math. Statist.*, **39**, 1711-1718.
- James, A.T. (1973). The variance information manifold and the functions on it. In *Multivariate Analysis-III* (ed. P.R.Krishnaiah). Academic Press, New York.
- Johnson, N.L. and Kotz, S. (1972). *Distributions in statistics: continuous multivariate distributions*. Wiley, New York.
- Kates, L.K. (1980). Zonal polynomials. Ph.D. dissertation, Princeton University.
- Khatri, C.G. (1966). On certain distribution problems based on positive definite quadratic

- functions in normal vectors. *Ann. Math. Statist.*, **37**, 468-479.
- Kushner, H.B. and Meisner, M. (1980). Eigenfunctions of expected value operators in the Wishart distribution. *Ann. Statist.*, **8**, 977-988.
- Kushner, H.B., Lebow, A., and Meisner, M. (1981). Eigenfunctions of expected value operators in the Wishart distribution, II. *J. Multivariate analysis*, **11**, 418-433.
- Macdonald, I.G. (1979). *Symmetric functions and Hall polynomials*. Oxford Mathematical Monographs, Clarendon Press, Oxford.
- Mardia, K.V., Kent, J.T., and Bibby, J.M. (1979). *Multivariate Analysis*, Academic Press, New York.
- Marshall, A.W., and Olkin, I. (1979). *Inequalities: theory of majorization and its applications*. Academic Press, New York.
- Mirsky, L. (1955). *An introduction to linear algebra*. Oxford University Press.
- Muirhead, R.J. (1982). *Aspects of multivariate statistical theory*. Wiley, New York.
- Nachbin, L. (1965). *The Haar integral*. Litton Education Pub., Inc.
- Olkin, I. and Perlman, M.D. (1980). Unbiasedness of invariant tests for MANOVA and other multivariate problems. *Ann. Statist.*, **8**, 1326-1341.
- Parkhurst, A.M., and James, A.T. (1974). Zonal polynomials of order 1 through 12. in *Selected tables in mathematical statistics*, Vol.2. I.M.S. publications.
- Saw, J.G. (1970). The multivariate linear hypothesis with nonnormal errors and a classical setting for the structure of inference in a special case. *Biometrika*, **57**, 531-535.
- Saw, J.G. (1977). Zonal polynomials: an alternative approach. *J. Multivariate Analysis*, **7**, 461-467.
- Srivastava, M.S., and Khatri, C.G. (1979). *An introduction to multivariate statistics*. North Holland, New York.
- Stewart, G.W. (1973). *Introduction to matrix computations*. Academic Press, New York.
- Subrahmaniam, K. (1976). Recent trends in multivariate normal theory: on the zonal polynomials and other functions of matrix argument. *Sankhya A*, **38**, Part 3, 221-

258.

Sugiura, N. (1973). Derivatives of the characteristic root of a symmetric or a Hermitian matrix with two applications in multivariate analysis. *Communications in Statistics*, **1**, 393-417.

Sugiura, N., and Fujikoshi, Y. (1969). Asymptotic expansions of the non-null distributions of the likelihood ratio criteria for multivariate linear hypothesis and independence. *Ann. Math. Statist.*, **40**, 942-952.

Towber, J. (1981). Personal communication.

## TECHNICAL REPORTS

U.S. ARMY RESEARCH OFFICE - CONTRACT DAAG29-82-K-0156

1. "Maximum Likelihood Estimators and Likelihood Ratio Criteria for Multivariate Elliptically Contoured Distributions," T.W. Anderson and Kai-Tai Fang, September 1982.
2. "A Review and Some Extensions of Takemura's Generalizations of Cochran's Theorem," George P.H. Styan, September 1982.
3. "Some Further Applications of Finite Difference Operators," Kai-Tai Fang, September 1982.
4. "Rank Additivity and Matrix Polynomials," George P.H. Styan and Akimichi Takemura, September 1982.
5. "The Problem of Selecting a Given Number of Representative Points in a Normal Population and a Generalized Mills' Ratio," Kai-Tai Fang and Shu-Dong He, October 1982.
6. "Tensor Analysis of ANOVA Decomposition," Akimichi Takemura, November 1982.
7. "A Statistical Approach to Zonal Polynomials," Akimichi Takemura, January 1983.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 7	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle)  A Statistical Approach to Zonal Polynomials		5. TYPE OF REPORT & PERIOD COVERED  Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s)  Akimichi Takemura		8. CONTRACT OR GRANT NUMBER(s)  DAAG 29-82-K-0156
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics - Sequoia Hall Stanford University Stanford, California 94305		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS  P-19065-M
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office Post Office Box 12211 Research Triangle Park, NC 27709		12. REPORT DATE January 1983
		13. NUMBER OF PAGES 95
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report)  UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report)  Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES  The view, opinions, and/or findings contained in this report are those of the author and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)  Zonal polynomial, homogeneous symmetric polynomial, uniform orthogonal matrix, eigenfunction, complex zonal polynomial, Schur function.		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number)  See reverse side.		

## 20. Abstract.

Zonal polynomials form one of the essential tools for expressing noncentral distributions arising in multivariate analysis. However, they have not been used very often or usually taught mainly because (i) the theory has been based on some branches of advanced mathematics, and (ii) the computational difficulty. In this dissertation a self-contained theory of zonal polynomials is developed in the framework of standard multivariate analysis. Zonal polynomials will be defined as characteristic vectors of a linear transformation in the vector space of homogeneous symmetric polynomials. From this definition almost all known properties of zonal polynomials are derived, as well as some new properties. In addition to theoretical considerations, computational aspects of zonal polynomials are discussed extensively.